```
   1 AAGCGATAGC TGAGTGCGGC GGCTGCTGAT TGTGTTCTAG GGGACGGAGT
  51 AGGGGAAGAC GTTTGCTCTC CCGGAACAGC CTATCTCATT CCTTTCTTTC
 101 GATTACCCGT GGCGCGGAGA GTCAGGGCGG CGGCTGCGGC AGCAAGGGCG
 151 GCGGTGGCGG CGGCGGCAGC TGCAGTGACA TGTCCAGCAT GAATCCCGAA
 201 TATGATTATT TATTCAAGTT ACTTCTGATT GGCGACTCAG GGGTTGGAAA
 251 GTCTTGCCTT CTTCTTAGGT TTGCAGATGA TACATATACA GAAAGCTACA
 301 TCAGCACAAT TGGTGTGGAT TTCAAAATAA GAACTATAGA GTTAGACGGG
 351 AAAACAATCA AGCTTCAAAT AGAGTCCTTC AATAATGTTA AACAGTGGCT
 401 GCAGGAAATA GATCGTTATG CCAGTGAAAA TGTCAACAAA TTGTTGGTAG
 451 GGAACAAATG TGATCTGACC ACAAAGAAAG TAGTAGACTA CACAACAGCG
 501 AAGGAATTTG CTGATTCCCT TGGAATTCCG TTTTTGGAAA CCAGTGCTAA
 551 GAATGCAACG AATGTAGAAC AGTCTTTCAT GACGATGGCA GCTGAGATTA
 601 AAAAGCGAAT GGGTCCCGGA GCAACAGCTG GTGGTGCTGA GAAGTCCAAT
 651 GTTAAAATTC AGAGCACTCC AGTCAAGCAG TCAGGTGGAG GTTGCTGCTA
 701 AAATTTGCCT CCATCCTTTT CTCACAGCAA TGAATTTGCA ATCTGAACCC
 751 AAGTGAAAAA ACAAAATTGC CTGAATTGTA CTGTATGTAG CTGCACTACA
 801 ACAGATTCTT ACCGTCTCCA CAAAGGTCAG AGATTGTAAA TGGTCAATAC
 851 TGACTTTTTT TTTATTCCCT TGACTCAAGA CAGCTAACTT CATTTTCAGA
 901 ACTGTTTTAA ACCTTTGTGT GCTGGTTTAT AAAATAATGT GTGTAATCCT
 951 TGTTGCTTTC CTGATACCAG ACTGTTTCCC GTGGTTGGTT AGAATATATT
1001 TTGTTTTGAT GTTTATATTG GCATGTTTAG ATGTCAGGTT TAGTCTTCTG
1051 AAGATGAAGT TCAGCCATTT TGTATCAAAC AGCACAAGCA GTGTCTGTCA
1101 CTTTCCATGC ATAAAGTTTA GTGAGATGTT ATATGTAAGA TCTGATTTGC
1151 TAGTTCTTCC TTGTAGAGTT ATAAATGGAA AGATTACACT ATCTGATTAA
1201 TAGTTTCTTC ATACTCTGCA TATAATTTGT GGCTGCAGAA TATTGTAATT
1251 TGTTGCACAC TATGTAACAA AACAACTGAA GATATGTTTA ATAAATATTG
1301 TACTTATTGG AAGTAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
1351 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
1401 AAAAA  (SEQ ID NO:1)
```

FEATURES:

```
5'UTR:          1-179
Start Codon: 180
Stop Codon:  699
3'UTR:          702
```

Homologous proteins:

Top 10 BLAST Hits

```
                                                                Score   E
CRA|108000024647144 /altid=gi|12728868 /def=ref|XP_002675.2| RA...   372   e-102
CRA|18000004923424 /altid=gi|4758988 /def=ref|NP_004152.1| RAB1...   332   5e-90
CRA|18000004937406 /altid=gi|131787 /def=sp|P05711|RB1A_RAT RAS...   328   1e-88
CRA|18000004952860 /altid=gi|131785 /def=sp|P22125|RAB1_DISOM R...   320   3e-86
CRA|18000004995539 /altid=gi|103720 /def=pir||D38625 GTP-bindin...   313   3e-84
CRA|18000004967528 /altid=gi|92339 /def=pir||S06147 GTP-binding...   297   2e-79
CRA|18000004880958 /altid=gi|464524 /def=sp|Q05974|RAB1_LYMST R...   282   9e-75
CRA|18000004908714 /altid=gi|466171 /def=sp|P33723|YPT1_NEUCR G...   253   3e-66
CRA|18000005175724 /altid=gi|7497231 /def=pir||T33781 hypotheti...   253   4e-66
CRA|335001098696672 /altid=gi|11558649 /def=emb|CAC17833.1| (AJ...   251   2e-65
```

FIGURE 1, page 1 of 2

BLAST dbEST hits:

```
                                                          Score    E
gi|12867866 /dataset=dbest /taxon=960...                   654   0.0
gi|12097820 /dataset=dbest /taxon=96...                    654   0.0
gi|12793758 /dataset=dbest /taxon=960...                   624   e-177
gi|12338056 /dataset=dbest /taxon=96...                    622   e-176
gi|11977068 /dataset=dbest /taxon=96...                    609   e-172
gi|10339840 /dataset=dbest /taxon=960...                   517   e-145
gi|10349761 /dataset=dbest /taxon=960...                   436   e-120
gi|10997958 /dataset=dbest /taxon=96...                    385   e-105
gi|10996533 /dataset=dbest /taxon=96...                    381   e-103
```

EXPRESSION INFORMATION FOR MODULATORY USE:
library source:
From BLAST dbEST hits:
gi|12867866 Fetal brain
gi|12097820 Adrenal gland
gi|12793758 Brain neoroblastoma cell line
gi|12338056 Adrenal gland
gi|11977068 Skin melanotic melanoma
gi|10339840 Uterus leiomyosarcoma
gi|10349761 Skin melanotic melanoma
gi|10997958 Placenta
gi|10996533 Placenta

From tissue screening panels:
Whole brain

FIGURE 1, page 2 of 2

```
  1 MSSMNPEYDY LFKLLLIGDS GVGKSCLLLR FADDTYTESY ISTIGVDFKI
 51 RTIELDGKTI KLQIESFNNV KQWLQEIDRY ASENVNKLLV GNKCDLTTKK
101 VVDYTTAKEF ADSLGIPFLE TSAKNATNVE QSFMTMAAEI KKRMGPGATA
151 GGAEKSNVKI QSTPVKQSGG GCC  (SEQ ID NO:2)
```

FEATURES:
Functional domains and key regions:
[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

       125-128 NATN

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 5
    1   59-61 TIK
    2   97-99 TTK
    3   98-100 TKK
    4 106-108 TAK
    5 122-124 SAK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 3
    1   35-38 TYTE
    2 106-109 TAKE
    3 127-130 TNVE

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

       30-36 RFADDTY

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 3
    1   21-26 GVGKSC
    2 147-152 GATAGG
    3 152-157 GAEKSN

[6] PDOC00017 PS00017 ATP_GTP_A
ATP/GTP-binding site motif A (P-loop)

      18-25 GDSGVGKS

[7] PDOC00579 PS00675 SIGMA54_INTERACT_1
Sigma-54 interaction domain ATP-binding region A signature

      14-27 LLLIGDSGVGKSCL

BLAST Alignment to Top Hit:

```
>CRA|108000024647144 /altid=gi|12728868 /def=ref|XP_002675.2| RAB1,
        member RAS oncogene family [Homo sapiens] /org=Homo
        sapiens /taxon=9606 /dataset=nraa /length=222
        Length = 222

 Score =  372 bits (944), Expect = e-102
 Identities = 190/222 (85%), Positives = 190/222 (85%), Gaps = 32/222 (14%)
 Frame = +3


Query: 129 GGCGSKGGGGGGGSCSDMSSMNPEYDYLFKLLLIGDSGVGKSCLLLRFADDTYTESYIST 308
           GGCGSKGGGGGGGSCSDMSSMNPEYDYLFKLLLIGDSGVGKSCLLLRFADDTYTESYIST
Sbjct: 1   GGCGSKGGGGGGGSCSDMSSMNPEYDYLFKLLLIGDSGVGKSCLLLRFADDTYTESYIST 60


Query: 309 IGVDFKIRTIELDGKTIKLQI-------------------------------ESFNNVK 392
           IGVDFKIRTIELDGKTIKLQI                               ESFNNVK
Sbjct: 61  IGVDFKIRTIELDGKTIKLQIWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESFNNVK 120


Query: 393 QWLQEIDRYASENVNKLLVGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 572
           QWLQEIDRYASENVNKLLVGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ
Sbjct: 121 QWLQEIDRYASENVNKLLVGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 180


Query: 573 SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC 698
           SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC
Sbjct: 181 SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC 222   (SEQ ID NO:4)
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	256.4	7.7e-75	2
CE00060	CE00060 rab_ras_like	170.0	3.9e-47	2
PF00634	BRCA2 repeat.	9.9	0.39	1
PF00056	lactate/malate dehydrogenase	3.9	3.4	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF00056	1/1	13	29	..	1	18	[.	3.9	3.4
CE00060	1/2	8	64	..	20	77	..	86.8	8.9e-23
PF00071	1/2	13	64	..	1	52	[.	111.9	4.8e-32
PF00634	1/1	57	79	..	13	35	.]	9.9	0.39
CE00060	2/2	65	140	..	110	188	..	81.2	2.9e-21
PF00071	2/2	65	173	.]	85	198	.]	142.4	4.5e-41

FIGURE 2, page 2 of 2

```
   1 TTTTGGGTGT GTGTGTGTGT GTGTGTGTGT GTGCCTTTAC TAGTGACTCA
  51 GGTCACAGTT TTCTGAGATT TTTTTTCTCC CCTCAAGACA GAATCTTGCT
 101 CTGTCGCCCA GGCTGGAGTG CAGTGGCCTC TCGGCCCACT GTAGCCTCCG
 151 CCTCCCGGGT TCAAGCAATT TTCCTGCCTC AGCCTCCCGA GTAGCTGGGA
 201 TTACAGGCAC GCGCCACCAT GCCTGGCTAA TTTTTGTATT TTTAGTAGAG
 251 ACAGTGTTTC ACCATGTTGG CCAGGCTGGT CTTGAATTCC TGACCTCGTG
 301 ATCTGTCCGT TTTGGCCTCT CAAATTCCTG AGATTACAGG CATGAGCCAC
 351 CGAGCCTGGC CAGTTTTCTG AGTTTTTATT TGAAATCAAA ATAAGCTTTT
 401 TTTTTTTTTT TAATGGGCTT TAGAGTCCAG GGTAACGAAC ACTTTTTGGT
 451 GCCTATTACT GAACCATTCA GGGTATTCCT GGGGTGGTGA CCGTGTTCAT
 501 TTCAGAAACC AACATGTTCA TTTCAGAAAC CAAACTCGGG TAACTTTTGA
 551 TAAGTTCATC AACTAAGGCC CATGGCAGAA TTTGAGGGCT AAGGGGTGTA
 601 ATTAGTGTAT GGGTAGAAAT AAGTGCCTTC TTTCTATATT TTGGCGTTGT
 651 AGGAATTTAA AGTGATTCTG CAGTAAGTCT CAGGAGACAA TTTTCTTAGT
 701 TCTTAGAAGT TGGAAGATAA ACTTTGGACA ATGTATTACA CTATGCCCTT
 751 TGTAATTAAA TAACTCAAGA TAATGTGTTA AAGTTTAGCG GAGATTTAAA
 801 TTCCTGAGCT GATTAAAGAG AGCTGTTAAG GCCATAGGTT TTTTAAAAAT
 851 GAGTTAATAT TACTCCCAGA AATTGTAGGC ACTATATAGT GATGAATTGC
 901 ATATTTTTAT TGCTTATTAT TTTCCAGTCT TGCAGAATGG CTCAGGGTTA
 951 GTAGCAACTA AAAGATAATA CATTACAATT CAACCTGAAG GCCGGGACGA
1001 AGGTAGGAAT TGGATTTTAG GCTGGCTCTG GGCTGTGTCC CTCCCATCCA
1051 TGGGATGTGG AGCCATTGAA GGTTGTGGGG TCACGATGCA GGTGCTGTCT
1101 CAGAAAGATA CATCCGACTG TGTGTGCAAA TGGGCTGGGG CGGAGAAGAG
1151 AGAGAGAGGT AGAGTCCATT TGGAGACTAC TGCAATAGCC AGGCTGACGA
1201 GTTAAGAGCG GGGCACAGTA AGAATGGGAA GAAATCTAAG AAGAAAATGG
1251 TAGTGCGCGG GGCCAACAAT GGACGATGAC CGAACCCAGG TGGGGATGGG
1301 TGAGTGACGA GAAGAACCGC TCCGTGCCGT CCAGGGAGCC CCTTGACTTC
1351 CCTTCTGTTC TTAGAGCGGA CGTCCTCCTA CCAGCCCCCA ACCAGCGCCA
1401 CCAGGGTGGC GCAAGCCTCA AGCTGGTCAG GTCAGCAACA GCCGCAACGG
1451 AGGCAGGAGC CGACACGCTC GTACCCCGGC CCCCTCCCCG CCCCCGCACC
1501 CCCGGCAGTC CCTCCGGTTT GACCACTCCC CCCGGTCCCT TGCCTCCCCC
1551 GACCCCCAGC CTCCGTCGGC CGCCGGCACC ACCCTCCGCC CCTCTCCGCC
1601 CCCTCCCCCG TGGGGCGCTG ACTCGCCCGG CTGCCACGTC TCACTGATGA
1651 CATCACTAGG GCAGCTCGGC CTTAGCCAAT CCGCCAGGGG GAGTCCGAGC
1701 GAAGTCCTAG CCAGCGAGTC AGAGGGGAGG GGAGCAGGGA GGGGCCGAGG
1751 GTGGGGAGGT GAGGGAGTGG GGAATGGGGC GGGCGACAAC CCTTCAGGTA
1801 CGCATGCCCC AGAGGCGCGG CGCTTGGCGG GAAGCTGAGT CCTGGCCTTG
1851 CGTCGCACTG TCTGTCCTCA GCTCGCGTAG CCGCGCTCGC GACTCCCTTT
1901 CCCGGCATGC CAGGCGGTGC GGCCGCCCTC TGGGCCGTGT AAAGGCCCCT
1951 CGGTCTAAGG CTTCCCTATT TCCTGGTTCG CCGGCGGCCA TTTTGGGTGG
2001 AAGCGATAGC TGAGTGGCGG CGGCTGCTGA TTGTGTTCTA GGGGACGGAG
2051 TAGGGGAAGA CGTTTGCTCT CCCGGAACAG CCTATCTCAT TCCTTTCTTT
2101 CGATTACCCG TGGCGCGGAG AGTCAGGGCG GCGGCTGCGG CAGCAAGGGC
2151 GGCGGTGGCG GCGGCGGCAG CTGCAGTGAC ATGTCCAGCA TGAATCCCGA
2201 ATAGTGAGTT CAGGAGAGCA CCGGTCGGCT GGGTCCGTGG GCCAGCTTGG
2251 GGGATCTTAA AGGGGTCGAG GAGGGTTGGG GCAGAAGTCG GGGCATCGGC
2301 TGGGGTGAGG CGAGGGTGAT GGGTCAGGAG AGGCTGGCGG CCGGGAGTCG
2351 GGCCCCATTG TCTGACGCGG AGGGGCGGCC GCGCGGGGGA GGGGTCGGGC
2401 CGGAGGGGTG AGCCGCCCGG GCCTGGACCG GGTCAGGTTA GAGGGCCTGA
2451 CTGCGGGGCG GGTGCTGAGG AAGCCTGCCG AGGGGCCTGG GGCGGTGTGA
2501 AGGGGTATCT TCTCTCGGAG GCAGTGACTT TTGAAGGAGG ACTTGTCTCT
2551 AAGGGGAGGG GATGGGGTGG GAGAGCCCTT CTAGAGGGCA CTGTCAGACC
2601 CTGCGCCCGC ACTCTGCGGA GCTGTCAGGA TCTTCGGGGT AGAAACCAGC
2651 TTTACTTGTA AATCCTGAGC TTGTTGGGTC TCTCTCCTTC CATCCTCCCC
2701 GCCAGGTTTC AGGTAATATG GATGCTTTTC GGGACTGCGT GGGATTGAGG
2751 GGAATGAGTA GATGGTGAGA AGCAACTGAA CATTTATTAG TTCTCTTTTT
2801 GAGTTGTGTC TTGGAGGAGT TGTTTAAGAG CTCGCCGGGT CCATTGCCCT
2851 CCTATAAAAA CCTGGGCATT TGTGAGAATT TTGTTTTTTT TTTTTTTAAA
2901 GAGGACACCT AAGTCATTTT GTCTTCTGTG GGTCAAGGGA AAAAAAAAAA
2951 ACTAAAGCCA AGAAATGTCT TTTTGATACT CGCAGATTAA AGGAAGCTTG
3001 CTGTCAAGTT GAAAGAGAAA CGAACGGGAC CTATGATAGA TCTGTATGTA
3051 GGTTTTGGAT TACCTGCTTG GATGCTTGCA GATAGGGAAT GAGGTTCCAT
3101 GACGTGTCAT GAAAAGTTAA TGCATTTCTT TTTCTTGCTT ACTCAAGAAG
```

```
3151 TCACCACAGC AGATGTGACA CACCTGGCAC CTTTCCTGGG AACTGGTGTT
3201 CACTTCCCTT GGGTAGAGTT TGTTGGGCTC TCCTCAATGG CCCTTTAAAA
3251 ATTTCCTCTA CAGTTTACAT GCATGTAAAG TAATGAATAA TTGGAAGAGA
3301 CCGAATTGGT ATTCCTTTTC AGTGTCAAAG GCCTTTGAGG GATGGGGGAA
3351 AATCAGTATT TGTTGTAAAA GTTGAGTTTA TTTGCTGGTT TGGTCAATTA
3401 CTGCTAGACA TTTTCCCCTA AAAGGTCCAC CCACCAGTTT AGCTGACTGT
3451 CATATGTGTG TCACATGGCT CTTGCAAAAT GCTTACAAGT TTTGTAATAG
3501 TGTGGCTTGA AGCTGAAATC TTTTGCACTA AACAGAAACC GTAGTATTTT
3551 ATTAGAATTT CATGCTTTAG AAGTTGAGGG TAGTGTTCTT GTAGTGACAT
3601 TTGCTGTGTT GACAGTTTAA AAAAATTTTT TTTTCAAGGG CTCCAAGGAC
3651 AAAGTTGGTT TTGCACAGTT GAACGGAGGT GAACTTGAGG TTCTTAATTT
3701 AGTAGTTTTC TTGGTAACAA TAAAGAACAT GGATTTACTG CTTTATCGAG
3751 GTTTATAGAC CTCTACTGTT CAGGAAATTT TCTGAATTTG CTATATATAT
3801 GTTTATTAGT GTAAATAAAT CTTCAAGATT AGTTGAGAAC TTTGACAAGT
3851 TACTCAGCCT CTGAATTTTT TTTCCCTTTT GTAAAATAGG ATAATTGGAG
3901 TCATTATTCC TGTCAGGGTA GTGGTGAAAT TCAAATGTAT ATAAAAGAAT
3951 TTGAAAAACT GTGTGAGCAT TCTTCAGGTG GTATGCATCA TTTTCATGAA
4001 AGGCATTCTA TTAGTACCAG GATTTAGGAA TATAATCCTT GCGCTTAAGA
4051 AGTTTAGATA TAGGCCAGGC GCGGTGGCTC ACCTCAGTAA TCCCAGCACT
4101 TTGGGAGGCC GAGGCGGGCG GATCCCGAGG TCAGGAGATC GAGACCATCC
4151 TCGGTAACAC GGTGAAACCC CGTCTCTACT AAAAATGCAA AAAAATTAGC
4201 CGGGCGTGGT GGTGGGCACC TGTAGTCCCA GCTACTCGAG AGGCTGAGGC
4251 AGGAGAATGG CGTGATCCCG GGAGGTGGAG CTTGCAGTGA ACCAAGATCT
4301 GGCCACTGCA CTCCAGCCTG GACGACAGAG CAAGACTCCG TCTCAAAAAA
4351 AAAATTATTT ATTGTTTTGA GACGGAGTTT CAATCTTGTT GCCCAGGCTG
4401 GAGTGCAATG GCGCAAATCT CCTCTCACCG CCACCTCCGC CTCCTGGGTT
4451 CAAGTGATTC TCCTGCCTCA GATTCCCGAG AAGTTGGGAT TACAGGCATG
4501 TGCCACCACT CCCGGCTAAT TTTGTATTTT TGGTAGAGAC GGGGTTTCTC
4551 CATGTTGGTC AGGCTGGTCT CAAACTCCCG AAGTGATCCG CCCGCCTCAG
4601 CTTCCCAAAG TGTTGGGATT ACAGGCGTGA GCCACCGCGC CCGGCAGAAA
4651 TAGATTTTAT ACATGTCAAA TACCAGTAGA TATAGCAAAT TCCAGATGTG
4701 TGGCATGGAT GAGAGCAACA AGATTTCAGG GGGATGGTGG GTTGTGGTTG
4751 GCTATCTGGG TTTTGGAAGA CTTTATAGAA GAGAGACCTG AAAGGGATTT
4801 ATCAGCAATT AGATTTGGAG GAACAGAGGG AGTGACTAGG AATTTTCAAG
4851 GGGGAGAAGA AGGAGGAATG GCTCATAAAT GACAAGGACA GTAATAAGTA
4901 AATACGGTGT CAAATCATCC TTTCTTTTGA AGACTAATGA CCTCAAAGGG
4951 ATCAAACCCA GAAACAGTTT TTATATTTTT TCTGGGATCA AATACATGGG
5001 TATCTGGCCT ACTATATTTG TATTCTAGAC TGTTTAGTAA AATAATACAG
5051 GAATTTGAGA AAACCTTTGC AAAAGTGTTA GTGAAAATTA CTTAGGGTGA
5101 GAGGAAGTGA GGGATATTTT ATTAGGGGAG GTCACAAGGG CAGTGAGCAA
5151 TCAGATTTTT AGTAATCTGA CTTAAGCAGT TTCTTTTTGT TTTAATGAAG
5201 CTTGTTATCT TTATAAAAGT AATTAGAGAA AATTTGGAAA ATAAAGGAAA
5251 GAAAGAAAG TTCTTTAGTG TTTTATCACG CAAATACAAG CTCATTCGTT
5301 TTTAACATCT TGTTCCAAAC TCCAAAGTCT TGCTTTCTCT TCAATTAAAA
5351 CTTTAATGGG TGGATGCTTT TCCTGCTTCC AGTATGTTAT CTTAATAACT
5401 AACAATGGTA TATTAGCTAA TGTTTACAAA TGTACTCCAG ATGTTCCTTA
5451 AGTTACTTTG GTTTATCATT ACCAATTTAT ATTGTTTCTT TTAGAAATTT
5501 ATAATCTTTG TTAATGGGTT CTGCTAAATT TGGTAGTGAA AATGGGATCT
5551 TGAGAAAAAA GATTCTGAAG CAACAGAATT TTTAGATTTA TATTGGTTTA
5601 CATAAGAGTT GGTAGCTGTA TTACTTTTTT TGTTTGTTTT GTTTTTTTTT
5651 TGAGACGGAA TCTTGCTCTG TCGCCCAGGC CTTGGCCTCC CAAAGTGTTG
5701 GGATTACAGG CGTGAGCCAC TGTGCCTGGC TGTTTGTGTT TTTTTTTGTT
5751 TTTGTTTTCT TTTCTTTTTC TTTTTTTCGA GATGGAGTCT CACTCTGTCA
5801 CCCAGGCTGG AGTGCAGTGG CGCGATCTTG GCTCACTGCA ATCTCTGCCT
5851 CCTGGGTTCA AGCGATTTTC CTGCCTTGGT CTCCTGAGTA GCTGGGATTA
5901 CAGGCATTTG CCACCATAAC CAGCTAATTT TTGTATAGAG TACCCAGCCA
5951 TCTCTAATGT TGATCAGGCT GAAGCAGGTG GATCACCTAA GGTCAGGAGT
6001 TCAAGACCAG CCTGGCCAAT ATGGCAAAAC CCTATCTCTA CTAATACAGA
6051 AAATTATCTG GGTGTGTTGG CTGGCGCCTG TAATCCCAGC TACTCGGGAG
6101 GCTGAGGCAG GACAATCTCT TGAACCTCGG AGGTGGAGGT TGCAGTGAGC
6151 CGAGATCACA CCATTGCACT CCAGCCTGGG CAACAGAGCA AGACTTGTCT
6201 CAAAAAAAAA AAAAAAAAA AAAAAAGGC AATTGAAAGT GTAATCTGAA
6251 CAGTTAAAAA AGTAGATAGA AAGGGTTAAA GCTTTTTTTT GAGGATCTGA
```

FIGURE 3, page 2 of 21

```
6301 AGAAAAATGT GGATTTTTTT TGAGCTACGT TTTGAAGCAG GCAGTGATTA
6351 TTTCAGCACA TTAAGAAATG CTTAACATGG CCAGGCGCAG TGGCTCACGC
6401 CTGTAATTCT CAGCACTTTG GGAGGCCGAG GTGGGCGGAT CATTTGAGGT
6451 CATGACCAGC CTGGCCAACA TGATGAGACA CTGCCTCTAC TAAAAATACA
6501 AAAATTAGCT GGGTGTGGTG GTGCACGCCT GTAATTCCAG CTACTCAGGA
6551 ACCTGAGGCA GGAGAGTCAC TTGAACCTGG GAGGCGGAGG CTGCAGTGAG
6601 TCCAGATCAT GCCACTGCAC TCCAGCCTGA GGGACAGAGT GAGACTCCTC
6651 AAAAAAAAAA AAAAAAAAAG AAAGAAATAC TTAACATTAT TCTCGTGATT
6701 ATTCTCATAA CATTTTTCAT AATCCACTGG CTTCCAGTGG ATTTTTTTAG
6751 TGTCAAGAAA ATAATTTTGA TTGGTTCATC TTTAAGGAAT GTGTTAAGAA
6801 TAAAGCATGT CTACCTGTCT TCAGTATACC AGCTAACTAT AGTAGGAAGA
6851 AATATAGTAG TCTACTTAGA TCAACTATAA TTCTTTAATG CAGAAAAAGT
6901 TTAAAGTATT TACCTTATTT TTAGCCCCCA TCCCCTTAAG TATATCATGG
6951 CTCCAGAATC TCTGAAAATG TTATCAGTCT TTCAGACTTT GCTCTTCTTT
7001 CATGTTATAC TCAAGAAACA TTTGACCTTT TTTTTTTTTT TTTTGCTTGC
7051 ATTGTGTTTC AAATAATTTT TAACAAAACT TAAGTGTTTG AAAGTGAAAG
7101 CAGGTTGTCT TTGTGACTTT TGGTGGTGGT TTGAAAAACT CAGAAAAGTT
7151 TAAAGAAGAA AGATAACTAG TATTCTCATT GTCCAGAATA TGATTTTTTA
7201 AATGTCTATA GAATATCACC ATCTGTAATT CTTCCGGTAA TTTAAGTATT
7251 CAGTAGTTGT ATAAAACCTT TAAAATATAT ATATTGAGAA TTTTGTGTGA
7301 ATGAGATGAT GAGATAATCT TGTAGGATCA TTTAAAGATA AGAACTGAGG
7351 CCTGGCACAG TGGCTCATGC CTATAATCAC AGCACTTTGG GAGGCCCAGG
7401 CGGTAGATCA CCTGAGGTCA GGAGTTTGAG ACCAGCCTGG CCAACATGGC
7451 AAAACCCTGT CTCTACTAAG CATAGAAAAA TTAATTGGGT GTGGTCGTGC
7501 CTGCGTGTAG TCCCAGCTGC TTGGGAAGCT GAGGCGGGAG AATCTCTTGA
7551 ACCCTGGAGG TGGGCATTGC AGTGAGCTGA GATTGCGCCA CTGCACTCCA
7601 GCCTGGGCGA CAGAGCAAGA CTCTGTCTCA AAATAAAGTA AAATAAAATG
7651 AAGATAACAA CTGAAATTTC ACATTAAAAA TTTTTTTGTA GCGACTGTGC
7701 CTCCTATGTT GTGCAGGCTG GTCTCAAACT CCTGGCCTCA AGCGATCCTT
7751 CCAAAGCACT GGGTGGGCCA CCATGTCCAG CCTGAAATTT TGCATTAAAA
7801 AATTTCCCGC TTTTGGCTGG GCGAGGTGTC TCACGCCTGT AATAGCAGTT
7851 TGGGAGGCCG AGGCAGGCAG ATCACTTGAG GTCAGTTCTA GACCGGCCTG
7901 GCCAATGTGG TGAAACCCTG CCTCTACTAA AAACACCAAA TTAGCTAGGC
7951 GTGGTGGTGT GCGCTTGTAG TCCCAAGCTA CTGAGGAGGC TGAGACAAGA
8001 GAATCGCTTG AATCTGGGAA AAAGAGGTTG CCGTGAGCCA AGATTGGCCA
8051 CTGCACTCCA GCCTGGGTGA CAGAGTGAGA TTCTGTCTCA AAAAAATAAA
8101 AAATAAAAAT TTCCCCCTTT AATCAAATTA AGTTAAAATG AGGGATGTTA
8151 GACAGTTTTT AACCATCAAA TATTTTAGTT TAGTTTTTTT TTTTTAACGT
8201 TGTCTTAAAG ATGGAAGTGC TTCAAAATCA AATCTTCCTT GCCAGTTCTC
8251 TACTTGGCTT CTTTTTTTTT CTTTTTGAGA TAGAGTCTCA CTTTGTCACT
8301 GGAGTGCGTT GGCGTGATCT CGGCTCACTG CAACCTCCGC CTTCCAGGTT
8351 TAAGTGATTC TTCCACCTCA GCCTCTCAAG TAGCTGGGAG TACAGGTGTG
8401 TGCCACCACA CCCGGCTAAT TTTTGTAGTT TTAGTAGAGA CAGGGTTTCA
8451 CTATGTTGGC CAGGCTGGCC TCAAACTCCT GACCTCGTGA TCCACCCACC
8501 TCAGCCAAAT TGCTGGGATT ACTTGTGTGA GCCACGCGCC TGGCTTCTAC
8551 TTGGCTTTTA AAGGGAATTT TGCTTTCTGA GTAATTTTAT TTCTCAGGTA
8601 TCTTGGTCTT TTTAATTCTG GAAGCAATCT TAATAATTTA TGTATGTGCC
8651 CTGTAATCCC AGCACTTTGG GAGGCCGAGG TGGGCGAATC ACGAGGTCAG
8701 GAGATCGAGA CCATCCTGGC TAACACGGTG AAACCCCATC TACTAAAAAT
8751 ACAAAAAATT AGCTGGGCGT GGTGGCAGGC GCCTGTAGTC CCAGCTACTT
8801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
```

FIGURE 3, page 3 of 21

```
 9451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9851 NNNNNNNNNN NNCCAGGCTG GAGTGCAGTG GCACAATCTT GGCTTACTGC
 9901 AACCTCTGTC TCCCGGGTTC CAGCATTTCT TCTGCCTCAG CCTCCTGAGT
 9951 AACTGGGACT ACAGGCGTCC ACCACCACGG CCAGCTAATT TTTATATTAG
10001 TAGAGATGGG GTTTCACCAT GTTGGCCAGG CTGGTCTCCA ACTCCTGACC
10051 TCAGGTGATC CGCCTGCCTT GGTCTCCCAA AGTGCTAGGA TTACAGGCGT
10101 GAGCCACTAC GTTTGGCTGC TTATCAGCTT TTTACCACTT TGTCGCCACT
10151 ACATTTTGGA ATTTTCCTTT GAGAATTAGG CAAAATGCCC AGACTCCCCC
10201 CCGGCCCCCG CTTTAGAGGG AGAGGGGAGC AATTAGACTA TTCCTTTGTT
10251 TCCCTATAGA AGGTGGGGCT GAGATTACTG CTTTGATATC TGGAATGTAA
10301 TTTAGGGAAG AAAATTTAGG TCTTGGCCTT TCTTTGGAAC CACCCTGGGA
10351 GTGTTGCAGA TTATTAATAG GGTAATGGTG GAATGATATT CAGGGGAAAA
10401 ATGGTCCTGA GGAGCCAGAG AACTAAGTGT TAGTTTGTTG GCTGACTGAA
10451 ACATGTGAGA GATAGGGTAC AGAAGAAGTA GGAAATAGTT TTCCTTGGTA
10501 CTTCTGTGAC AGGTTGGCTC AATTGGCTGG AACACCCTAC ACTGCTTTAT
10551 TAAATCCAAG GTTGTGATAG GTTCCAGTTA AGTTACTGT GTTCTATGCT
10601 TGTAGATTTC CTAATTAGGA CAAGTAGTGT TAAAATATGCA TGCCTTTATT
10651 CACAAGAGGG ACCATTCTTT TGGAAACATC ACTTTTTAAT AATACTAGGT
10701 GCTATTTAGC ACTTACTCGG TGCCAGCCAC GTGGCTATGG TTTTTTTTTT
10751 TTTTTTTTTT CGAGACATGA TCTAGCTCTG TCTCCCAGGC TGGAGTGGTG
10801 GTAGCACAGT CATGGCTCAC TGCAGTCTCA ACCTCCTGTA CTCTAGTGAT
10851 CCTCCTGTCT CAGCCTCCTG AGTAACTGGC ACCATGCCTG GCTAATTTTT
10901 TTTAAGAGAT GAGATGTCGC TATGTTGCCT ATGCTGGTCT CGAACACCTG
10951 GGCTCAAGTG ATCCTCCCCG CCTGAGCCTC TCAAAGTGTT GGGATTACAG
11001 GTGTGACCCA CCTCACTTGG CCATCTATGG TCTTTACATA GGGCATTTTG
11051 TGCAGTCTGC ATCTCAAACT AGTGATCTTC AACAGTGAAA CTCAGTGAAT
11101 TATGTAATTC ATGTTTTCCA AGAACAATGA TGGATTTAAT TTCTCTGAAT
11151 GTATTTCCTT TGTATAATAA TAGTACTTAA GTGGAATTAC TCTTTGTCCT
11201 TTCTACTCTC CTTATAGATA TTTTCTGGTA TCTTGATTTG GGACTGTTAC
11251 ATTTAACCCA TTTATGGTCG TGTAGCCATA CTCACGTTAC ATTTGATGCA
11301 TCTGCTCCCT TTGTGTCTAT ATACTCATAT AACATTTTGC ATAAAGTTAT
11351 AGGCAGTTCA CACCAAGGCT GTTCATGAAC CTCAGATTAA GAATACTTGA
11401 TTTAGGAGAT TGAAAACAGA AAGAGAATG TTAACTATCA TTATCAATAT
11451 TAAAATGTGA AAATCTGAGA GTGACAAAGC TTAGCTTTAA ATCTGGTATC
11501 CCAAACTCAT TTGAGTTTTT TTTTTTTTTT TTTTTTTTTT GAGACAAGGT
11551 GTCGCTTTGT CCCCCAGGCT GGAGTGTAGT GGTGTGATCT TGGCTCACTG
11601 CAACCTCCAC CTCCCAGGTT CAAGTGATTC TCCTGCCTCA GCCTCTGAAG
11651 TTGCTGGGAT TACAGGCTGC GCCACCACGC CCAGCTAATT TTTTGTATTT
11701 ATAGTAAAGA CGGAGTTTCA CCTTATTGGC CAGGCTGGTC TCAAACTCCT
11751 GATCTTGTGA TCCTCCCGCC TCGGCCTCCC AAAGTGCTGG GATTACAGGT
11801 GTGAGCCACT GTTCCCGGCC TAATTTGAGT TTTAAAATGT GGAGTTTAAG
11851 ATGTTAGTCT TAAAGTGGGT TAGATGAAAT TTATAAAAAT AGTCAAATAG
11901 CTAAATTTAT AAAAGGCCAT TTGAAACAAT TTTGTGAAAT ATATAATGTG
11951 GATAATTATG TAGTGCTTTA TGTGTAGATT GGTGGTTAGC ATCTGCCTGA
12001 TGAAGAGCAG TTGGATTTCT TACTTACTAA AGCTAGTGAA ATCTGAACTC
12051 CAAATTAGGC ATCTTCACCA GGCTTTTTTG AGCCGAGCTA ACTTACTCTC
12101 TTTTTTATTT TTATTTTTTA ATTAATTAAT TTTTTTTTTT TTTTTTTTTT
12151 TTTGGTAGAG ACAGGATCTC CCCATGTTAC CCAGGCTTGT CTCTGGCTCC
12201 TTGGCTCAAG CAGTCCTCCT ACCTTAGCCT CCCAAAGTGC TAGGATTACA
12251 GCTGTGAGCC ACTGCGCCAG GCTGAGCTTA TTCTCTACTA ACACAAGTGT
12301 TCTAATTTAA TTTAAGCAGT GAATCACACT TTTCTTTGTA TTTGGTCAGG
12351 TTCTGGGTGC TAGTTTATAT ATGATTTGAT TCATTCTGAT AGGGTTTTTT
12401 TGTTTTTTTT TGTTTTTGTT TTTTTGTTTT TTTTGAGACA GAGTCTAGCT
12451 CTGTCGCCCA GGCTGGAGTG TGGTGGCTCG ATTTCGGGTC ATTGCAACTT
12501 CTGCCTCCCA CCCAGGCTGG AGTGCAGTGG CTCGATTTCG GGTCATTGCA
12551 ACCTCTGCCT CCCAGGTTCA AGCGATTCTC CTGCCTCAGC CTCCTGAGTA
```

```
12601 GCTGGGATTA CAAGCACCCA CCACCATGCC CGGCTAATTT TGTGTATTTT
12651 TAGTAGAGAC TGGGTTTCAC CATGTTGACC ACGCTGGTCT CGAACTCCTG
12701 ACCTCAGGTG ATCTGCCTGC CTTGGCCTCC CAAAGTGCTG GGATTACAGG
12751 TGTGAGCCAT CACACCAGGC CTCAAGAACT TTTTATTTTT GAGACAGGGT
12801 CTCACTCTGT CACCCAGGCT GGAGTACAGT GGTGAGATCA TGGCTTACTG
12851 CAGCCTGGAC TTCCCAGGCT CTGGTGATCC TCCCATCTCA GCCCCTGGAG
12901 TAATTAGGAA TATAGACACA CACCCATGCC TGGCAGTTTT TGTATTTTTT
12951 TTCTTTTTTC TCTTTTTTTG TAGAGACTGG GTTTCACATG TTGTATCAGG
13001 CTGGTTTTGA ACTCCTGAGC TCAAGCAATC CTCACTCTTT GACCTCCCAA
13051 CGTGCTGGGA TTACAGGCAT GAGCCACTGT ACCTGGCCTT TTCTACATTA
13101 AAAACTTTTT ATTAAAAAAC CCAAATCTTC CTTGTGGTTG TATATACATA
13151 TATACATAGG TACACACATG GAGAATTTTA CCTTGGAGGA AGGCTTGGTA
13201 AAGAAAATAG CCCTTTGGGC CGGGTGCGGG GGCTGACGCC TGTAGTCCTA
13251 GCACTTTGGG AGGCTGAGGT GGGCGGATTG CCTGAGCTCA GGAGTTCAAG
13301 ACCAGCCTGG GCAACACAGT GAAACCCTGT CTCTACTAAA ATACAAAAAA
13351 TCAGCTGGGT GTGGCAGCAT GTGCCTGTAG TCCCAGCTAC TTGGGAGCCT
13401 GAGGCAGGAG AACTGCTTGA ACCCGGGAGG CAGAGGTTGC AGTGAGCCGA
13451 GATTGTGCTA CTGCACTTCA GCCTGCGCGA CAGAGCAAAA CTCTGTCTCA
13501 AAAAAACAAA CAAACAAACA AAAAGGAAA ATAGCCTTTC TCTATCATCA
13551 GAGTATATTA AGAGTTGAGT TTTTTTTTCT GTTTTTTAAA ATTTTTGTTG
13601 TTTATTTTAA ATTACAAAAC ATGGACTCTG CTTACAAATT AAGAAAATGA
13651 CTCATGTTCA AACAAGCATA ATCAATATAA CAGTTAATAC AAGTTAAATA
13701 TTGTAATATG TTTACGGAAT AGCATGGCAA AATAGTGCAA AAGATTTGGG
13751 GAAGGGGCCT ATAATTTCTG TTAACAGAAA GTTTTAGTTA TGTTGATTCA
13801 ACTGGAGAGG AACAGAGCTC CCAGAAGGAC TCCAGAACAC TTGATGCTTG
13851 TCTGAGTGGG GTCAGCAGCA CTGAGTTCCC ACCAGCCAGA AAGTTTGTGT
13901 GTGTACATTA TTTCCCTTAA CTGCCACAAT AATCCCATGA AGAAAATGCC
13951 CTAGTTTTAC AAACAAGGAA ACAGAGGCAG AGAGAGAGTTA AATGACTTGC
14001 CCAAGGGCAT TCAAAGTAAG CAACTGAATT GGAATTTTAA CTCAAAGGCT
14051 TGGATGTCCC ACTACAACAA ATAGGCTGTT TCTGCTTTAC TACATGTGCT
14101 TACTTCTAAG AATTTAACAT TTTAGGCTGG TTGTGGTGGC TCACTCCTGT
14151 AATCTCAGCA CTTTCGGAGG CTGAGGTGGG TAAATCACTT GAGCTCAGGA
14201 GTTTGAGACC AACCTGGGCA ACATGGTAAA ACCTCATCTC TACCAAAAAA
14251 AAAAAAAAAA CTAGCTGGAC GTGGTGGCAC GCGCCTGTGG TCCCAGCTAC
14301 TCAGGAGGCT GAAGTAGGAG GATCGTTTGA GCCTGGGAGG TGGAGGTTGC
14351 AGTGAGCCCA CATTGCATCA CTGCACTCTA GCCTAGGTGA CAGAGTGAGA
14401 GCCTATCTCA CACACAAAAA AAAGAATTTA AAATTTTAGT CAAGTAATTA
14451 GGCACTAACA TTTTGTGGTC AGTTACTTTA CGAATTCATG GTTGGAGGCC
14501 TGATGTGGTG GCTCATGCCT GTAATCCCAG CACTTTGGGA GGCTGAGGCA
14551 GGAGGATTGC TTAAGGCCAA GAGTTCAAAT CAGCCTGAGC AACCTAGTAA
14601 GATCCCCTTT CTGCAAAAAA TTTAAAAATT AGCTGGGCAT GGTAGTGTGC
14651 ACCTGTAGTC CCAACCACTT GGGAGGCTGA GGTGGGAGGA TTGCCTGAGG
14701 CCAGGAGTTT GAGACCTGGG CAGCATATGA AGACCCTGTC TCTAAAAAAC
14751 TAAAAATAAA AAATAGCCAG GTGTGGTTGG TGTGCTTGTG GTCCCAGCTA
14801 CTCAAGAGGC TGAGGCAAGA GGGTTGCTTG AGCCCAGAAG TTGGAGGCTG
14851 CCGTGAACTG TGATTGCACC ACTGCACTTC AGCCTGGGTG ACATAGCAAG
14901 ACCCTGTCTC TGTGGTGGTG GTGGGTGGGG GTGGGGGAAG GGATTTAAGA
14951 AGGGTTTGTG AGGTATGTAT TATTTATAAA TGGGCTTTTA ACTTTACCCT
15001 TCACATCTTG GGTTGAAATT AATTGTATCC ATTCTCAGTT TTTCTGTCTT
15051 GCTATATATT TAAACTTGGA GACTTAGAGG TCATGGATGT CTTTCTATGA
15101 AAAGCAAATG AAGCAGAGGG CTGCCTTCTC TTGCTGTAGA GGGCACACTT
15151 GCTGCAGAGC ATGTTACTGT TTTATGCATT GCTAGGCTTT GGGAGTTGTG
15201 ACTTGTATGA TCATAGTACT TACAACTATT AGTTGGCAAT TTTTAAACTT
15251 TAACTTTAGA TTATATATGT AAACTCCTGT GTTCCTTTGT CACTGATAAT
15301 CTGAACAGAA GCCTTGGATA AATAATTTTG AAGTTTTTGT CTGAACCTCT
15351 GAAATTTGTA TTGTTATCTC ATGGTTTTGC TGGGAGGAAG GAGAAATAAC
15401 AATGGCCACT TACTGTGCTT CTGTATGTGC CAGACAGTAT GTGCTAGATG
15451 TTTCAGAAAC GTGATTTGTA ATCCTGACAA GAAGCCTAAT TGGGTGGTAG
15501 TGGGTGCTAA TTGAACCTTA TAGATGAGGA AATTGAGGCT CATGGTGGTA
15551 AGTGAATAAC TTGCACCAAG ATCCTATGGC TGGTATGCAG TAGAGCCTCA
15601 ATTCAAGTAC GGGTCTTCCA GGTCCAAACC CATGCAGGCT TTGAGAGGTA
15651 AGGAGGTAGA GAACGTTGAC ACCCCCTTCT TGGTGTGTTT TTCAGCAAAT
15701 ACTTGTATGC ATATTAAAGA CTGTCTACCC TTTTGTCATC TTGTGTCACT
```

FIGURE 3, page 5 of 21

```
15751 TGCTGCTTCC TTTGGTACTA CCCAAATTTC TTTCAGCATT TCAGCTTTGA
15801 ATTTTTATTT TTATTTTATT TAATTTATTT ATTTTTTTGA GATGGAGTCT
15851 CACTCTGTTG TCCAGGCTGG AGTGCAGTGG CGTGATATCA GCTCACTGCA
15901 ACCTCTGCCT CACAGGTTCA AGCAATTCTT CCTGCCTCAG CCTCCTTAGT
15951 AGCTGGGACT GGAGGTGCCC ACCACCACGC CCAACTAATT TTTGTATTTT
16001 TAGTAGAGAT AGGGTTTTAC CTTGTTGGCC AGGCTGGTTT TGAACTCTTG
16051 GCCTCAAGTG ATCCACCCAC CTCGGCCTCC CAAAATGCTG GGATTACAGG
16101 CATGAGCCAC TGCACCTGGC CAGCTTTGAA TTTTTAGAAT ACTGTTCTAA
16151 ACAGAACTAT ATTGGAACCT GGAAAATTAA TCTATTGTCT CTAAATACCA
16201 AAGAAAAACA TGTAATTTTA GTGGTTGATT ATGGGAACAA TTTTTTTTAA
16251 GATGGTTCAT CTGAATGGGA AGCATTTTTT TTTTAATTGC TTGACTATTT
16301 CTTTAAATTT GGAGAAAAGA CCATTGCCCT CTCAGATTTC TGGTAATTGG
16351 TCACATTGAT CATTTATATT GACTGACAGG CTGCTTTGTC CACAGCTGAA
16401 GGATTGTTTA ATTTTTTTTA AATTATAAGA GTAATATGTG CTCACTGTAA
16451 AATTCACAGT ACAGAAGCAT ATGAACTAAC TAAAAGTTCT TACCTCTTGT
16501 CTCCAGCAAG GAGTAAGTGT TTCAACCTGA AGGTTGGTTT TGAATTGTGT
16551 TCTGTGGAGC GTACTTAAAG TGAGTGAAGA AGAAAAATTT ATGTCAATCA
16601 TGATCATTGC AGCTGAAGTT TTTATTGTTT CACCCCCTAA AGGTTATTAA
16651 AATAGTATGT AGTTTAGTAG TCTTGATAAT TTTCCCTTAA GATTTATTGG
16701 CCAGTATATC AGGATTTTGT TTTAAATTTG ATATGTGAGC TTAGTTTTAT
16751 GCTATTTTCA AATAAGACAT TTAGAAGAAG ATAAAATAAC ATTCCTGTCT
16801 TAGTCTGTTT TCTGCTGCTA TAACAGAATA GCACAGACTG GGTAATTTAT
16851 AAACAGTAGA AGTTTATTTG GCCTGTGGTT CTGGAGGCTG GGAACTTCAA
16901 GAGCATGGTT CTGCCCTTTG TGCTGTGTTA TCATATGGTG GAAGGTGGAA
16951 AGGCAAGTGG GTATGTCAAG ACAGAGAGCA AGAAGGGGCT TGAACTCACT
17001 TTTATAACAG AGTGACTCCA GAGATAGCTA ACCCACTTTT GAGAGAATGC
17051 ATTAATCCAT TCATGAGGGC AGAGCCCTTG TGACCTAATC ACCTCTCATT
17101 AGGCTCTGCA TCCTTAAACT GGTTTTTTTT TGTTTTTTTT TTTTGAGACG
17151 GAGTCTCGCT CTGTTGCCCA GGCCGGACTG CGGACTGCAG TGGCGCAATC
17201 TCGGCTCACT GCAAGCTCCG CCTCCCGGGT TCACGCCATT CTCCTGCCTC
17251 AGCCTCCCGA GTAGCTGGGA CTACAGGCGC CCGCCACCGT GCCCGGCTAA
17301 TTTTTTGTAT TTTTTTAGTA GAGACGGGGT TTCACCTTGT TAGCCAGGAT
17351 GGTCTCGATC TCCTGACCTC ATGATCCACC CGCCTCGGCC TCCCAAAGTG
17401 CTGGGATTAC AGGCGTGAGC CACCGCGCCC GGCCCCCCTT AAACTGTTGT
17451 ATTGGGGATT AAGTATCTAA CACAGGAACT TTGGAGGATA CATTTAAACC
17501 ATAAGAATTC CTGTCATGCA AATGAATCCA TTCTAGATGA AAGAGAATGA
17551 ATTTAGTTTC CATTGAACTT TATAAATAGG CCTTTTCTAA GGTACTTACA
17601 GCTGATATTA TAAAATTTAT ATTTGTTTTT ATAAATTTGT ATTTGTATTT
17651 CTGTTTGTAC AAATACAATT ATACACTATA GTTCTCTGCT GTTAGATTTT
17701 TTTTCTTCCT TAGCATGTTT CCAAAGGGTG GAATGTTGAA AGTTGGGTTA
17751 ATGTCAATCA GCTTTCTTTT GTAAAGTGTT CATTGACATG TGAACCTTGT
17801 CTGAGAATCT AAATTTTATT TCATGAAAGA AGAAAACAGT ATATTCTCAT
17851 TTAACCCAGA ATTTAACTTC ATATACTTGT GGCTGTATTG GGAGTATGCC
17901 ATTGCTGTCT GTTTACAACC TGACCTACTC TACCTACTTA GAAGTAATTT
17951 GTGTTATGAT AGGTGTGCTG TGCTGACATA TGCTGAACAT ATTTGTAAGG
18001 GTGTTAAGTC ATTGAATAAA ACGCTTTTCT CCTCCTTTCA AATAACATTT
18051 TTTATTTCTG GTTATAAAAG TCATACAAGC TTACTGCAGG TTGTTAAAAA
18101 GGTATAAAGA AGAAACCGTC AATCCATTAT AATCCTACAG TTTAGACTTC
18151 CTGCTCCAGC CTCTCAGAGT GCTGAGATGA GCTAGCCATG CCCAGCCCCT
18201 CAAAAGATTT TTTAAAAAAC AAAAATGAGG TTATACTTTA AAAAATTCTA
18251 TATTCCTTTC ACATAACAGT GTTATTTTGG AGGTTTTAGA ATTTCCAGTA
18301 GCATTTTAGA TTCAGAAACA AGCTGATTCA TCCTCTACTT TGTACTTTAG
18351 GCAAGAAAAG AATTTTACCT AAATAGAATT TTGAACTGAA AATCTGTTTT
18401 TCTAACTTTT TATTTAAAGA ATATTGTTCC ATGCTTTCAC AGTAGTGACT
18451 TTTAATTTTT ATATTTTTTA TTTTATTTAT TTAGAGATGG GGGTCTCACT
18501 CTTGTTGCCT AGGCTAGAGT GAGTGCAATG GTTCTATTCC TAGCTCACTG
18551 CAACCTTGAA CTCCTGGGCT CAAGTTACCC TCCTGCCTCA GCCTTCTAAG
18601 TAGCTGGGAC TACAGGTGTG CACCACTGCA CCAGGCTTTT TTTAAAGGCA
18651 TAGAAAATGG TAGTGCTTGC ATACAAAAAT GGCGTAGGTA CATACATCAG
18701 CGGACATCAA GACTATGTTC AGATCATAAA TGTACATATA TGTACCGATG
18751 CCATTTTTGC ACGCAAACAA ATAATGGAAA TTGAACTCTA AACTGAAATT
18801 TGAAACAAGG GTTCTGGGGT GGGCCCTCTT GCTGATTTGT AATTGAATGT
18851 ATAGTTCAAT TTTTCCCCAT CTGTTAAGCA AAAGACAATT CTAATGTTAG
```

```
18901 CAAAAATCCA CATATCCTGT CATTGATCAT TTTTTCCTTA ATTTTCTTTA
18951 AGAGATGGGG CTTCTCTCTA TGTTGCCCAG GCTGGTCTGG AACTCTTGGG
19001 CTCAAATGAT CCTCCAGCCT CAGCCTCCCA AAGTGCTGGA ATTAATAGGC
19051 ACAAGCTGCT GTGCCTGGCC CTGTCATCAG TCATTTAACT TCATGCAAAC
19101 TGAGTAGAAT AAAACTCGTC CTTACTGTAC CTTATTGCTT TTGTTTTATT
19151 GTTGGAACCT CCAATATTGC GAAAGTAGAC CAAAAGTTGA CTTATAGGAA
19201 AAACTGATAG CAAAAATAAT TTTTCTCTTG TTGCTGTATT TCATGCCCAC
19251 CATCCAGTTG TTAAAGCCTA CTGTTAATTT CTCTCAGCCT CCTCCTTTCT
19301 GTCCAGGCTT ATTCTATGCC ATTCTTACCT TAACTGTTTT TAGCTTTCTC
19351 ATAGAGTGAA CTTTTTAAAT TAAAATAAAA TATCTGCTCG TAGTATTATA
19401 AAATTCAAGC AGTTCAACAG AATTTTTCAC TAATAGAAAT ACTTGTACCT
19451 CAAAAGCAGC TTTATTTTAC AAACCCAGCC CAATTTGTGA TTAGATTTAA
19501 CTTGAGAAAA CATGAAATGT CTCTCATATT GTTTAAAAAT ATCATAAGTG
19551 GCTGGGCACG GTGGCTTATG CCTATAATCC CAACACTTTG GGAGGCTGAG
19601 GCAGGTGGAT CACTTGAGGT CAGGAGTTTG AGACCAGCCA GGNNNNNNNNN
19651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19951 NNNNNNNNNN NNNNNNNNNN NNNNNNNTTC ACCATGTTGG CCAGGCTGGT
20001 CTCAAACTCC TGACCTCAGG TGATCCACCT GCCTGGGCCT CCCAAAGTGC
20051 TGGGATTATA GGCTTGAGCC TCGCCTGGCC TCCTCATAAT TTTTTAACCT
20101 TTATAAAAAC CTTTTCTAAA ACCCTTTTTA TTTTGAACTA AATTTAGATT
20151 TACTGAAATT GTGAAATCAA TGTGGAGTTC TTGTATACCC TTCTTTCCGC
20201 TTTTCCTAAT AGTAACATCT TACATACATG GTACATTTGT CCAAATTAAG
20251 AAATAAACAT TGGTACAGTG TTAACTATAG ACTTAATCTG GTTTCTCTAA
20301 TTTTTTCACT AATGTTCTTT TTCTGTTCTA GGATCTAATT CAGTATACCA
20351 TATTGTATTT AGTTGTAGGC CATGTTAGCC ACCTTCAATC TGTGACAGTT
20401 TCTCAGTCTT TCCTTCTTTT TCGTTATCTT GACAAGTTTG AAGAGTGCTG
20451 ATAGGTATTT TATAGAATGT CCGTCAGTTG TCTGTCAGTT TGTATTTGTC
20501 TGATGTATTT TTTTTTTTTT TTTTGAGATG GTGTCTCGCT CTGTCGCCTA
20551 GGCTGGAGTG CAATGGCATG ATCTTGGCTC AATGCAGCCT CCACCTCCGG
20601 GGTTCAAGTG ACTGTCCTGC CTCAGTCTCC CAAGTAACTG AAACTACAGG
20651 CATGTGCCAC CACGCCTGGC TAATTTTTTG TATTTTAGTA GAGAAGCAGT
20701 TTCACCGTGT TGCCCAGGCT GGTCTCGTGC TCCTGAGCTC AGGCAATCCA
20751 CCCGCATTGG CCTCCCAAAG CGCTAGGATT ACAGGTGTGA GCCACCATGC
20801 CTGGCCAATA TTTTGAGGGA TATACTTTGG TGAGGTCATG CAGATATCCT
20851 GTTTCTCCTT AGTTTTATCG ATTAATTTAG CATTTATCCA GTAAATCTTC
20901 CTTGCAGCAA TTATTTTTTC TTTTTCTTTT TTCCTTAATT TTTTTTTTAA
20951 GAGATGGGAT CTCACTCTGT TGCCCAAGTT GGAATGCAGT AGTGAGTTCA
21001 TAGCTCACTG CAGCCTCAAA CTCCTGGGCT CAAGTGATCC TTCTGCCTCA
21051 GCCTCTCAAG TAGCTGGGAC TACAGGCATA GACCACCACA CCCAGCTAAT
21101 TAAAAAAAAT ATTTTTAGAG ATGGGGGTTT TGCTATGTTG CTCAGGCTGG
21151 TCTTGAACTT GCTGGCCTCA TGTGATCCTT CTACCTCAGC CTTACAAGTA
21201 GGTGGGAATT ACAGGTGTGA GCCACCACAC CCAGCATTGC AGCAATTATT
21251 AATGTAGTGC TACTGGTCAT TTTCTGTTTT TCTCATTTCT TCAGCATGTG
21301 TTATTGACTT GTCTCTTCCC TCCCATTTAT AATCATTTAT ACTGCTATGA
21351 ATTCATGAGT ATTTATTTTG TGAGTTATAA TCTAATACGT ACTTAATTTA
21401 TTTTGTGCCT CAAATTGTTC TGGCTTGGCC ATTTTTTTTT TTTTTTTTTG
21451 AGACGGTCTC GCTCTGCTGC CCAGGCTGGA GTGCAGTAGC GCCATCTCTT
21501 CTCACTGCAA CCTCCACCTC CCGGGTTCAA GCGATTCTCC TGCCTCAGCC
21551 TCCTGAGTAG CTGGGACTAC AGGCGTGTGC CGCCACACCC GTCTAATTTT
21601 TTGTATTTTT AGTAGAGACA GGGTTTCACC ATGTTAGCCA GGATGGTCTC
21651 GATCTCCTGA CCTCGTGATC TGCCCGCCTC AGCCTCCAAA AGTGCTGGGA
21701 TTACAGGTGT GAGCCACCAA GCCCGACCGG CTCCTGTATC CTTTTAACAT
21751 GAGGTGCTGT CATCATTTTT TCCCCCTAAT ATTTTGGCCA AAAATGTTAA
21801 TCAAGGATGG CACAAATTTT CTGTAGCTGT ATCTCACAAT GAAAGAGGCC
21851 TGATTAAAAA TGTAAAACTA AAATGTTCTC TGATCTCTTA GCACATGCTT
21901 TGTAAAAGGC ACAGTGCTAG ATCCTTGTAT ACGTAGGATGA GTAAGTCAGC
21951 TTACCTTCCA CACCCACAGA TAGCTATGTC AAACGTAAGG GTGGAGAAAC
22001 ACAGACCCCA AACTTCTCGA GGGTAGAAAA TATGAGGTTA TAGTAGATTA
```

FIGURE 3, page 7 of 21

```
22051 GAACTACAAA AAGCTAGAGG AAGTTCTGAA CTGGAAACAG TGGATAGGAT
22101 TTACTAGAAT AATTTACGAG GGTGACAATT GTAAATCTTC ATAGGTTTCT
22151 TTTTTTTCCT TTCTCTTTTT TTTTTTTTGA GATGGAGTCT CGCTCTGTTG
22201 CCCAGGCTGG AGTGCAATGG CGCAGTCTCT CCTCACTGCA ACCTCCGCCT
22251 CCTGGGTCCA GGTGATTCTC CTGCCTTAGC CACCCAAGTA GCTGGGATTA
22301 CAGGCATCTG CCACCATGCT GAGCTAATTT TTGTATTTTT TTTTTTAGTA
22351 GAGACGGGGT TTCACCATGT TGGTCAGGCT GGTCTTGAAC TCCTGACCTC
22401 AGGTAATCCA CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGGTGTGA
22451 GCCACCGCGC CCAGCCAAAT TTTTATTGGT TTCTAAACTA GCGTAATTTA
22501 GTTTTTTTCA CTTAAGTCAA AATTATATTA TTGTAGGATA AAAACTTAGT
22551 GATCCAAATT CATGAGGAAT GAAGAATAAA TACATTTAAA GTCTTACCAT
22601 TTGCTAAATT AGTCTTGGCT CTTTGTACCA AAATTCTGTC CTTGTGCTCT
22651 GTAATTTTAT ATTTGTATAT TTTCTATCAA CATTTTTACT GTGTGGTGTT
22701 TTGTAAATTA TAAAAACGTT TTAAAGCAAA CTCAGAACAA TGAATTCTCA
22751 CGAATATTCA GTATATTTAC AGTTGAGAAA TAAACTACTT CTGTAGTAGG
22801 TAATTTAAAA TGTCCCAATG CAAGTTAACG TGTCACTGAT CACGCTATTC
22851 AGGTGTGTGT CTTTGATAAG GGGAGGTGGG GAAGTTGTG GGTTTGATTT
22901 TATTTGCCTT TCTCATGTGA CTGTTGTCAT GTTAGTAAAC AAATGGTTTG
22951 CGAGAGAACC AGTAGTCTTT TGCAAAGATT GTCTTATACA GAGCACTCAA
23001 TTCTTCATAT TATTTATAAT GGCTTTAATT TAAGCCTTAA ATTATTAGAA
23051 ACTCATAAAT AATTTTTTTA TTTGTTTTTT TGAGATGGAG TTTCGCCCTT
23101 ATTGTCCAGG CTGAAGTACA ATGATGTGAT CTTGACTCAC TGCAACCTCC
23151 GCCTCTCGGG TTCAAGTGAT TCTCCTGCCT TTGCCTCCCA AGTAGCTGGG
23201 ATTACAGGCA TGCGCTACCA TGCCTGGCTA ATTTTGTATT TTTAGTAAAG
23251 ACAGGATTGC ACCATGTTGG CCAGGCTGGT CTCGAACTCC CAACCTCAGG
23301 TGATCCACCT GCTTCGGCCT CCCAGAGTGC TGGGATTACA GGCTCACTGA
23351 GCCACTGTGC CCAGCCATAA TGCGTTAAAA TAAGAGTGTT ATATTTGTAA
23401 AACTTAAAAA AATGTAGTGG TTGAAAAAGG TAATTTAAAA AGAATTGACT
23451 ATTAATTTCT TGAAACCATA ATGTAACTTG TAGTTCAATT AGGAAACCTT
23501 CATGTTTCTT TCTTTCTTTC TTTTTTTTTT TTTTTGAGAT GGAGTTTTGC
23551 TCTTGTTGCC TAGGCTGGAG TGTGTGATGT CAGCGCACTG CAACCTCTGC
23601 CTCCTGGGTT CAAGCAATTC TCCTGCCTCA GCCTCCCGAG TAGCTGGGAT
23651 TACAGGCGCC TGCCACCACA CCCAGCTAAT TTTTGTATTT TTAGTAGAGG
23701 CGGGGTTTCA TCGTGTTGGC CTGGCTGGTC TCGAACTCCT GACCTCAGGT
23751 GATCCACTGC ACCTGGCCCC CGTTCATGTC TTTTAAAGCT TTATGGTTGC
23801 TCTGAAATAG AGTTGTTGAT TTTTTTTTTT TTTTTGAGAC TCCTCTTTTG
23851 CCCGTGCTGG AGTGCAGTGG TGTGATCTGA GCTCACTGCA ACCTCCACCT
23901 CCTGAGTTCA AGCAATTCTC ATGGGTCAGC CTCTCAAGTA GCTGAGATTA
23951 AAGCTGCCCA CCACCATGCC TAGCTAATTT TAGTATTTTT AGTAGAGATG
24001 GGGTTTCACC GTATTGGCCA GGGTGGTCTG GAACTTCTGA CCTCAGGCAT
24051 GAGCCACTAC GCCTAGCCTG GGTTGTTGAT CTTTAAGGTG ATACTTCAGG
24101 CAACATCTGA GGCCCAGTAC AGTCCTTTAC TTCAACTGGC TCCAGTACAG
24151 CAAATTCAGG GAATGTTTTT GAGTGTTTAC TGGATGCCTG GCGTGGAGTT
24201 CAGGGAGATT GGTACATTGA GTCCAGTTGT TGTGTTGAAA CTTCTGTTTA
24251 AAAACCTCCC TACTAAGTCC CAGCTACTCA GGAGGCTGAG GCCTGAGAAT
24301 CACTTGAACA CCTGGAGGCA GAGGTTGCAG TGAATCGAGA TCGAGCCACT
24351 GCACTCCAGC CTGGGCGACA GAGTGAGACT GTCTAACAAC AAAAACAACA
24401 CCCCCCAAAA AACCAACCTA CTATGGTAGT ATCAATGCTG TGATAGTCTT
24451 CCTTTCTTCA TACAGGTAAA TTCTTAACAT ATACTCATTG TTAATGTTCA
24501 GTGTTCAGTA TTCTTAAGAG TATTTGGGGC CAGGCACGGT GGCTCATGCC
24551 TGTACTCCCA GCACTTTGGG AGGCTGAGGT GAGCAGATTA CCTGAGGTTA
24601 GGAGCTTGAG AACAGCCTCC AACATGATGA AACTCCCGTC TTTACTAGAA
24651 ATACAAAAAT TAGCTGGGTG TGTTAGCACA TGTCTGTAAT CCCAGCTACT
24701 TCAGAGGCTG AGGCAGGAGA ATTGCTTGAA CCTGGGAGGT GGAGGCTGCA
24751 GTGACCTGAG ATTGCTTCAC TGCACTCCAG CCTGGGCAAC AGAGCGAGAC
24801 TCTTGTCTCA AAACAAACAA ACAAAAAAAG AATATTTGGG GCCAGGCATG
24851 GTGGCTCACA CCTGTAGTCC CAGCACTTTG GGAGGCCAAG GTGGGTGGAT
24901 CACTTGAGAT CAGGAGTTGG AGACCAGCCC GACCAACATG GCTAAATCCC
24951 GTCTCTACTA AAAGTACAAA AATTAGCTTG AGCAACAGAG CAAGACTCTG
25001 TCTCAAAAAA AGAAAGAAGA ATATTTGGTT TAATTAAGAA GGAACCTTAT
25051 CAATAGTAGT AAAGTCAGCC AGCTGAACTG CCAAGTACAA ATTGTTGGTA
25101 TTAGGTATCA ATCATTTATT AAGGATAATA TTCTACAATA GCGATCTTTT
25151 TAAAAATTTT AAAATCTCAA ACTGGAAAGG ATGTCTAGTT CATTCTATGC
```

FIGURE 3, page 8 of 21

```
25201 TTCAGTCCCC TCTTCTGATT TACTTGTTTA GAAGATTTTT GTTTCCTTCT
25251 CTGACTTCTA TTTTGCTGCT GACTGGCACT TGGGATTTTT AAAAAATTAT
25301 TTTCCTCATA TATAATTAAA GACAATAAGT ATAACAATAA GTATAATATG
25351 GTAATTTGCT AAAACCCAAA CAATGTTTTA AGTAATGCAT ATCATTATGT
25401 AAACCTACGT AATAGTTGAA TATTCACAAA GATAATCGCT TATAGAAGTT
25451 TTATATCCTC TCTTCTTTGG CAGTGCAATT AAAACAAAAA AAATAAGTTT
25501 TATGTCTTGT TTACATGTAA ATAATTTTAA TCTAAATTGT GACGTGGTTT
25551 TCACTTTAGC ATATTTTTGA AAGTAAATCA AAAAGGACAA AATACAAAAT
25601 CATGTATATC TTCTACAAAA ACGATATATA AATTCTAAGG TTTTTGTCCT
25651 TTTGAAATTG CTTAAAAGAA TGCATAGAAC TGGTGTCTGA GTTGGGAAGG
25701 ATCTATGAGG GATTTCCTTG GAGACCGTGG GTGAATAATA ATGTTGTCTT
25751 AGTTCCATGA AGGAATCTCT GGGGATAGTT TTTGAGTTAG GCCTGGCAAT
25801 GTTAGAGATA CATAAAGAGA GCCTTGTTTT ATCACTGGGT GCGGTGGCTC
25851 ACACCTGTAA TTCCAGCACT TTGGGAGGCT GAGGCGGGCA GATCATGAGG
25901 TCAGGAGATC GAGACCATCC TGGCCAACAC GGTGAAACCC GTGTCTACTA
25951 AAAATACAAA AATTAGCTGG GCGTGGTGGC GCATGCCTAT AATCCCAGCT
26001 ACTCGGGAGG CTGAGGCAGG AGAATCACTT GAACCAGGGA GTTGGAGGTT
26051 GCAGTGAGCC GAGATCGCGC CACTGCACTC CAGCCTGGGT GACAGAGCAA
26101 GACTCCGTCT CAAAAAAAAA AAGCTTGGTT TTCAATGGTT CTGAAAAATG
26151 CTTTAATACA AGTGTAGAGT GTTAGTCAAG TTTTGCACTT GGATAAACAG
26201 CCTGTGAATT TATCACATTT CTAGTTTATA ATATGGGCTT TCAGAAGTTA
26251 TATGAACATT GTTTTGACGG GAGAATTCAA GCTGGATGCT AGAGAAGGAT
26301 CGTGAGAACC CCTTCATTGG AGGAGTGCTA TGAAATTATT TGATCTTGGA
26351 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTGAGAC AGAGTTTCGT
26401 TCTTATTGCC CAGGCTGGAG CTGGAATGCA GTGGCACGAT CTCGGCTCAC
26451 TGCAACCTCT GCCTCCTGGG TTCAAGCAAT TCTTCTGCCT CAGCCTACCA
26501 GGTAGCTGGG ATTACAGGCA TGCGCAACCA TGCCCAGCTA ATTTTTGTAT
26551 TTTTAATGGA GACGGGGTTT CACCATGTTG GTCAGGCTGG TCTTGAACTC
26601 CTGACCTCAA GTGAACTGCC TGCCTCAGCC TCCCAAAGTG TTGGGATTAC
26651 AGGTGTGAGC CACTGCGCCT GGCCTGATCT TAGAATTTGA AGGAGAGACT
26701 AATATTTCAT GGGCAAAAAC AATGAAAAGT TACCTTTCTG TATTCTAATA
26751 CTATAGAGGA GTGGGATTTA TTTAGAATGT TTTAAGTATC TTGGGCAGTC
26801 CAAGAGTGCG TATCACTTAT TTTTCTTTTC CTTCTTTCTT TTTAAGTGGA
26851 AGTTCACTGA TGTTAGAGAT CATAGGTGGC ATTGCCTACT TTTTACATAA
26901 TTTTATCATG TTTAGTGATC TGTCAGAAGG GCTGTGGCTG TTTGCAGTTT
26951 TGGCTTAAGC CATGCATGGG CTTTATAGGA GATGTAGTCT TCACAGTGAG
27001 TTGTTATTTG TAGCTGTGTT TTTGTTTTTG TATAGCTTAT AGCAATGCAG
27051 TGTGCTTTTT ATTAACATCA TTTTCTTTTT CTTTTTGCAG TGATTATTTA
27101 TTCAAGTTAC TTCTGATTGG CGACTCAGGG GTTGGAAAGT CTTGCCTTCT
27151 TCTTAGGTTT GCAGTAAGTT GAAATTGAAA TGTCTTTACA ATTAATGGTA
27201 CAATTAATGC TATGTATGTT TTCTAGGTAG ATAAAATTAA ACAGTTTTAT
27251 TCAGAATAAG TTAATTCTTC CAGAATTTAT ATATTTAAAG ACTCCAAATA
27301 TACATCCCCA GTGGTATCTT GGACTGTTAA ATAGAAAAAT ATTGTTGCTC
27351 TTAAAAGAAA TTCAGTGAAG TCTGGTTATA AAGTCAGAAT GTCTAATACT
27401 TTTGGTCAGA GTCAAACAGC AGTTCCAATA TAGGCAGCAA GTTAAAGGGG
27451 TAGTTGGTGG CCTGTGTTGA AAGCGACTTG ATGAAAATAA ATCTTTAAAT
27501 TAAACTTTAG TAGAATAAAA AGAAAAAGCA GAGCCAGGTG ACGCAGTGGA
27551 TCATGCCTGC AGTCTCAGCT ACTCAGGGTG CTGAGGGTGG AAGGATCACT
27601 TGAGTCTAGG AGTTTTGAGA CCAACCTGGA CAACATAGCA TGACTCTGTC
27651 TCTGAAAAAA AAAGTTAATA AAAGAAAAAG TAGGGTCTTG GACAAACTTC
27701 GTTGGCCAAT GGCATAGTTC TAAATGCTGA AGCTGACAGA TAAAGGACTT
27751 TTGACTTAAC AGAATCCACA GTGTCCTTCA TAGTCTTTAT CAACTACCTT
27801 TAAATTTAGC ATGTTTCCTG GCCAGGTGCG GTGGCTCACG CCTGTAATCC
27851 CAGCACTTTG GGAGGCCGAG ACGGGCGGAT CACAAGGTCA AGAGATTGAG
27901 ACCATCCTGG CTAACACGGT GAAACCCCGT CTCTACTAAA AATACAAAAA
27951 ATCAGCTGGG TGTGGTGCCA CACGCCTGTA GTCCCAGCTA CTCGGGAGGC
28001 TGAGGCAGGA GAATCGCTTG AACCCAGGAG GCGGAGGTTG CAGTGAGCTG
28051 AGATGGTGCC ACTGCACTCC AGCCTGGCAA CAGAGCAAGA CTGTCTCAAA
28101 AAAAAAGAA AAAAATAAA AAAACAAATT AGCATGTTTC CCTTCTAGAG
28151 ATCATTGTTT CTCAGAGCAT GGACCAAAGA CTCCTGGGGG TTACCAAGAC
28201 CCTCTCAGGT AGCCCATGAG GTCAAAATAT CCTAATAATA CTAAGATGTT
28251 AGTATTTGTA AGGAAATATT TACTTGGTAA TAATACTAAT ATAAAAGATG
28301 TTTGCGTTTT TCAGTGATGA CATTGGCTCT GGTACAAAAG CATGTGGGTA
```

```
28351 AAATTGCTGC TGGCTTGGTA CACATCAAGG CAGCGCTAAG CTCCAAATTG
28401 TACTCATGGT GATGGCATTC TTTACCTCTG TGCCCTCACA GGAACAAAAA
28451 CAAGCCGTGC CATTTTTATT GAAGATTGTC CTTGACAAAA CAGTTAAAAT
28501 GATTAATTTT TGAAAAATGT TGATCCATGA GTATTCCTTT AAAAATATTT
28551 GTGAAGAAAT GGGAAGTTCA CATAAAACAA TGTTTTTTTT TTGTTTTTTT
28601 TTTTTTTTTT TTTTGAGACA GATTCTGGCT GTGTTGCCAA GGCTAGAGTG
28651 CAGTGGCGTC TGGCTCCCAG GCTCAAGCTG TTCTCCCACT TCAGCCTCCC
28701 AAGTGGCTGG GACCTCCCAA GTGGATGCGC CATCATGCCT GGCTGATTTT
28751 TGTATTTTTT TGTAGTGACA AGGTCTCACT GTGTTGCACA GGCTGGTCTC
28801 AAACTTCTGA GCTCAAGCGA TGCATGTGCC TCAGCCTCCC AAAGTGCTGG
28851 AGAAAGCACT TTTTACTGCA TACTGGCTAG TGTGTTGGTT ATTTTGGAGA
28901 AAAGAAAAGC ATTTGTAGTT TTTTGAGTTG TAAGCTGAGC TAACTGCTTT
28951 ATTTTTTTCT GTGGAACACC ATTTCTTTTT TTTTTTTTGA GATGGAATAT
29001 TGCTTTGTTG CCCAGGCTGG AGTGCAGTGG CACAATCTCG GCTCACTGCA
29051 ACCTCCGCTT CTCGGGTTCA AGCAATTCTT CTGCCGTAGC CTCCCAAGTA
29101 GCTGGGATTA TAGGCACCTG CCACCAAGCC CAGCTAGTTT TTGTATTTTT
29151 AGTAGAGATG GGGTTTCACC ATGTTGGCCA GGCTGGTCTC GAACTCCTGA
29201 CTTCGTGATC CGCTTGTCTC AGCCTCCCAA AGTGCTGGGA TTACAGGCGT
29251 GAACTACTGC ACCTGGACAT TTTTTTTTTT TTTTTAACTT GAAAGAACAG
29301 CTAACAGACA GATTAGAACA GAATTGGCTA TTTGACAGAT TTTCTCAGAT
29351 GAACTGTGAT AGTCATTTCA AGGGAAGTAG CTGCAAGCAT TTGTTGGCTG
29401 AAATAAAATT TAAGTTTATC ATGGAAAATT AGAATTTGAA AAAACTTAGA
29451 GTTTACCACT TGACAGTATC CTAAATACAT ATGACTTTTC TGATGAGTGC
29501 CGATATTAAT GAAGGTTATT TAAAAAATAT TAAATAATGT ATAATTCTTT
29551 TTATATAACA GTTAAAAATA AAACCATGAG TACTAGAATA AAACATAGGT
29601 GGCTCTTTAA TCTTGGTTTG TGAAGGTATT TTTTAAAATA AGAAAAAAGC
29651 AAGAAATCAC TGCTAAATTT GACTATTAAA ATTAATTTAT CACAGGCACA
29701 AAAATGTTAG AAAACTAATG GCAATAGCAA ATATATATAT ATGAGGATTG
29751 GTATTCTCAA CATATAAAGC ACATTTGCAC ATCAACAAGA AAAGAATATT
29801 TCTCCTAATG GAAATAGTGG CAAATACATG AGCAGTCAGT TGAAAAAAGA
29851 AGTAATACAA ATTGCTGGCT GGGTGTGGGT GGGGTCACGC CTGTAATCCC
29901 AGCATTTAGA GGCTGAGGCT GGCGGATCAT CTGAGGTCAG GAGTTCGAGA
29951 CCAGCCTGAC CAACATGGAG AAACCCTGTC TCTACTAAAA ATACAAAATT
30001 AGCCGGATGT GGTGGCGCAT GCCTGTAATC CCAGCTACTT GGGAGGCTGA
30051 GGCAGGAGAA TTGCTTGAAC CCAGGAGGCG GAGGTTGTGG TGAGTCGAGA
30101 TCGCACCATT GCACTCCAGC CTGGGCAACA AGAGCGAAAC TCCATCTCAA
30151 AAAAAAAAAA AAAAAAAAAA AAAAGGAAGT AATACAAATT GCCAATAAAT
30201 ATGGAAAAAA AAAAAGGCTC AACTTTATTT GTAATTAAAG GCCTTTAAGT
30251 TAAACTTAGG TGTCATTTAA TTTTTATTAA ATTGGCAAAT ATTAAAATTA
30301 AGCATAATTC TTAAGCAACT CTCGGTAGGT GGGAAGAATC TAGCTGTAGC
30351 CTCAGGTGTT TGTGCCTCAA GGAAAACCCT CTCTGGGATG TCCATTGCTT
30401 GAAGTCAAAG GTTTTCCAAT AATACCTGGA AACTATTTTT AAAATGCTGA
30451 TCCCCATACC CTCAAAATAT TAATAGAGAC AATCGTGAGG ACTATAATAA
30501 AGAAATGTGC AATAAGCTCT GGGGGCACAG AGGGAAGAAT CTATTGGCTG
30551 AGGAGTTGAA GAAATTGTTT GGACACTCAG TATTGCCTGA GCTCAAAACT
30601 GAAGGATGAA TAAATGCCAC ATGACCTTGG GGCTGGGGAG TAAGTAGGGT
30651 TATGCAGAGA GAGATAACTG AGGCTTTTGG GCAGACGAAT AGTAACGGCT
30701 CAGGCATGGG AGTAAAGGTC ATTTAGAGAT TTACAAGAAT TCAGCATTTC
30751 TTTCTTTTTC TTTTTTTTTT TTGAGATGGA GTCTAGCTCT GTCATCCAGG
30801 CTGGAGTACA GTGGCATGAT CTCAGCTCAC TATAACTCCC ACCTCCCGGG
30851 TTCAAGTGAT TCTCATGCCT CAGCCTCCCG AGTAGCTGGT ATTACAGGCG
30901 TGTACTACTG TGCCTGGCTA ATTTTTGTAT TTTTAGTAGA GATGGGGTTT
30951 CACCATGTTG GTCAGGCTGG TCTCCAACTG CTGAGCTCAA GTGATATGTG
31001 CACCTCTGCT CCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGTACCC
31051 GGCCAAGAAT TCAGTATTTC TATCCAAGTA CCTGGGGGAT AGATGTGCTA
31101 CATGAATATT TATTGCATTC ATTTTGTTCT CTGCATTTTT TTTTTTTTTT
31151 TTGGTTTGAG ATGGAGTCTC GCTCTGTCGC CCAGGCTGGA GTGCAGTCGT
31201 GCAATCTCGG CTCACTGCAG CCTCCACCTC ATGGGTTCAA GCGATTCTCC
31251 ATCTTGGTCT CCTGACTAGC TAGGTTTACA GGCGTGTGCC ATCACACCCA
31301 CTAATTTTTT GTATTTTTAG TAGAGACAGG GTTTCACCAT GTTGGCCAGG
31351 CTGGTCTTGA ACTCCTGATC TAAAGTGAGC CTCCCACCTT GGCCTCCCAA
31401 AGTGCTGGGA TTACATATGT GAGCCACTGC GCCTGGCCTC TATATACTTC
31451 TATAGTACCT GATACTTATT AGGCACTCAA TTACAACATA ACTTTTTTTT
```

```
31501 TTTTTTTTTT TTTTGAGACA GAGACATGCC TTGTCGCCTG GGCTGGAGTG
31551 CAGTGGCACA GTCTCGGCTC ACTGCAACCT TCACCTCCCG GGTTCAAGTG
31601 ATTCTCCTTC CTCAGCCTCC CGGGTAGCTG GGATTACAGG CGCCCGCCAC
31651 CACGTCCAGC TAATTTTTTG TATTTTTAAT AGAGATGAGG TTTCACCATC
31701 TTGGCCAGGC TGATCTCAAA CTCCTGACCT TGTGATCCAC TCACCTTGGC
31751 CTCCCAAAGT GCTGGTATTA CAGGTGTGAG CCATCATGCC CGGCCCATAT
31801 TTCTAAAAAC ATTTTCTTAT AAAATGACAT TGCCATTATC AACCTGCAAA
31851 ATACATTTCC ATTTGGTTGT TTTCTTGCTT AGTCTTTTAA TCTAGAGTTT
31901 TATACCTTAT CTTTTTTATT TATATATTTT TTATGTCATT GACTTTTTGC
31951 AGAAACTGAA GCACTTGTCC TGTAGATTGT CCAATATTCT AGATTTGTCA
32001 TTTTGTTTCC TTGTGATGTC CTTATGCTTA TTTGTTTGTC CCTCTTTCTG
32051 TAATTAGAAG ACCTAGAACT GCACTATCCT TAGAGTAGCT ACTAGCTCTA
32101 TGTAGCTATT TAAATTTAAA TTAATTAAAA TTGAAAAAGT TTGGTGGCTC
32151 ACACCTGTAA TCCCAGCACT TTGGGAGGCC AAGGTGGGAG GATTGCTTGA
32201 GTGCAGGAGT TCAAGGCTTC AGTAAGCTAC GATTGTACTC TAGCCTGGGA
32251 GACATCAAGA CCCTGTCCCT TTAAGGGGGA AAAATAATTG AAAAAATCAA
32301 AAACTTAGTT TCCTTGTTTC ACAAGCTGCA TAGGGCTAAT GGCTACCATA
32351 TTGGCTAGCA CAGCTTATAG AACCTTTCCA TTGTCACAGA AAGTTCTGTT
32401 TGGCAGTGCC GTTCTCATTA GACCTGATTC GATTAAGGTC CATCTTTGTT
32451 GACAGAGTAC TTCTTAGGTG GTGCTTTGTG GTTCATATGA TGATAGCCTG
32501 GTCTGTTCAT TCATATATCT TTTCACGAGA AATATTTTTA TTCCATTCTG
32551 AATAAAATTT CATGGCAGGT ACTTGCAAGA AGCAGTTATA ATTTTAAAGT
32601 TTAACATTAG GTTAAAAAAT TGACAGGAAA CATATATTCA CAGGTAAAAC
32651 TTGTACACAA ATGTTCATGG CAGCATTATT CATAATAGCC AAGAAGTGGA
32701 AACAACCCAA ATCAATTTAT GAATGGATAA AATGTTGTAT ATTTGTAGTA
32751 CATGTAATAT TATTCAGCCA ATAAATGGG CCAGGCATGG TGGCTCACAC
32801 CTGTAATCCC AGCACTTTGA GAGGCTCAGG CAGGGGGATC ACTAGAGGTC
32851 AGGAGTTTGA GACCAGCCTG ACCATCATCA CGAAACCCTG TCTCTACTAA
32901 ACGTACAAAA ATTAGGCAGG CGTGGTGATG CACGCCTGTA GTCCCTACTA
32951 CTCAGGTGGC TGAGTCATGA GGATTGCTTG GACCCCGGGA GACAGAGGTT
33001 GCAGTGAGCT GAGATCATGA CACTGCACTC CAGCATGGGC AACAGAGCAA
33051 CATCCTGCCT CAAAAAAAAA AAAAAAAAA AAAAGAAGTA CTGTTACATG
33101 GTACAACATG GATGAACCTT GAAAACATTC TGCTAAATGA AGGAAGACAG
33151 ACACAGAGGG CCACATATTT TATGATTCCA TTTATACGAA ATGTCCAAAA
33201 TTGGCAAATC TAAAGAGAAA GTAGATTAGT GGTTGCCAGG GAGTGAAGAC
33251 GGGTTCTTTC TGGAGTGAAG AAAATGTCCT GGAATTCGTG GTTGTAGTTT
33301 GCAACCTTGT GAATGTATAA GGACCACTGA ATTGTCCACT TCAAAAGGGT
33351 GACTTTTATG TTATGTGCAT TATATCTAAA AAAAAAATCA TAATTAGGAA
33401 GCAAGATTGA CTTCTAAGAA AAAGCGGAGT GAAATTGTTG TTTTGTGGTG
33451 AATAAATTGG GTGGGTGGGT CGCAAGAGTT TTGCTGATTA GTGATTAGAA
33501 AAATTATTCA TAATCATTGA AAATATAAAA TATTTTTCTA TATGATGTAT
33551 GTAAAGAATT TGGCAAGAGA TGATGTTTGG AAAAAATAAA GAATGGCTAT
33601 TGTAGAGATC TTAAGGAAAG AAACTACAGT TAAGTAGTGC TTTGTAATCA
33651 GAATATGAAG TAAGTACTGA AAGTGGATGG AGTGGCTGTT GTCAGCATGT
33701 TATACTTTAT ACATTTCATT CATAAATTTG GACTGTAGAT AAAAGTAAAC
33751 TTTTTTTTTA TTTACTCTTG AACAACAGTT TTTTTTTTTC CACTTAGACT
33801 TGCATCTGCT CCACTGAACA ATACATTTAA TTGTTAATTA TTTCCCCCTT
33851 CAGGATGATA CATATACAGA AAGCTACATC AGCACAATTG GTGTGGATTT
33901 CAAAATAAGA ACTATAGAGT TAGACGGGAA AACAATCAAG CTTCAAATAG
33951 TAAGTGACTT GGCTAGTAAT TTTTTTGAAA TTTATTTTGG TAAATTTGTA
34001 ATGTATTGTT ATTTTGTATA TATTTACTAT GCTAACAAAA TTGAATGTAA
34051 AATGTCTTAA GATTCATGTA CTTAAGATAG AATGGTAGAA TAAGAATTAC
34101 TTAGATTAAA AATAATATTT TCAAGATTAC TTAAGCCTCA TTGAATTTTC
34151 TGTTCATGAA GCAGAGAAAC TCATGTTTTA AGTCAAACTT GGTCCTCATC
34201 TTTTTCTTTT ATCAGTGGAA ATCTAAGTTC AAGTTTACCT TGTCCTACAC
34251 TGCAAATGTT ATAGACCATT TTTGTTTGTC TTTTACTGTG CTAAGTGCAT
34301 GGAACATTAA AGGAACCCTA GGAAGAGATT CTTCATATGT GGCTCAGTTG
34351 AAGAGAAGTA CTTATGTAGT TCTAAGTATT TTTATTAGAT AGTGTGCACC
34401 AACTCTGTAG AAACACAGAA TTTTGTTGGA AAAAGGAACT TAGTTTTTGT
34451 AACATGTTCA TTTTACTGCT CAAAAAAACG AATGCTGAAA GATTTAATGA
34501 CTTGCCTACA GTTACTGGTA GAACCAAGTG ACCGAAGCTC TGTCTTCAAT
34551 ATTTTGTGTC TGTGTGCCAT CCTATCCCCC TTATCCATCT TTACACCCCC
34601 AGCCCCCAAT TAAATATAGG CAATTATAAT AGTTCAGTTG TGCCTCTTCA
```

```
34651 GTATGGGTCT GAGTCCTGTC AGTGTGGGCA TATCTGTGGT CTTTTAAAAA
34701 ATAAATCTCT CAGTATTTTT CAGAGTAGGC TATTAGCAAG AAGTAGGCTA
34751 TAAACACAGG AAACCAGTGA CTGCCCCTTT TCATGGAACT GATGACACAT
34801 GGAATTGGAA GGAGTCCTGC ATTAGGAGTC AGAAGACTTA GATTTGTTGT
34851 CTTGGTTCTA GTATTTACCT GTTAGAGAAT CATGGGTTTG TGTCTCTGGG
34901 GAAAAGGCCG AAGTAACCCT GAGACCCAGT TTCCTTTCTA AAATGTGTGT
34951 GATGACACCT GATTACTAA TTTATAAGCT AGTTGTGAGA ACCAACTGTA
35001 ATAGCTTTGT GTATGTGACA ATACGTGTGA AAGCCCTTTG TAAACTTTTG
35051 GGCAGCATAT AGATACTACT TATGATATGA CATGCCCAGA TAAATGGGTG
35101 TTTGATAGGT TAAGTTGCTC CCTTTTCTTA CATGACTCTG ATGAGGAAAA
35151 GAAGGTATGT TAACAAAAGA TAGGTGGCTG TGGATATTGA TATAAGTAAA
35201 CACACTTGAT GTGTCAAATT AGGACTTGCA AGGATTTAGT TTTCAGAAAT
35251 AGCTTGAAAT ACTTTCAATC AGTGAACAAA TTACCCTCCA TATTTTTTCC
35301 CACGATATAA GTACAGTCTC AACCTTTTAT TTGGCACCAT AAAGAGCACA
35351 TAAAGATCTA CCCAAAACTG TACTTTAAAG CACTGGTATG GAATAATTGT
35401 ATTATGTGTG ATCATTGGTG TTTATAAGAT TTGGGTGTGT ATTCGTGTGT
35451 GAAACATTCA TATTTTGTTA CTTTCCTGTG GCTGGAAGGG ATCTTATAGG
35501 ACACTGTCTT TCATCTTTGT CTGTCTTTCA TCTTTAATAG GAATTTCTTT
35551 TCCATGCCTG AAGGCCTCAT TTTGAACATT TTGTTTGTTT GTTTTTTTAT
35601 TTTTTGAGAT ACAGTATTGC TCTGTCTCCC AGGCTGGAGT GCAGTGGCGC
35651 GATTTGAGCT CACTGCAACC TCCGCCTCCT GGGTTCAAGT GATTCTCCTG
35701 CCTCAGCCTC CCTAATAGCT GGGATTACAT GTGTGTACCA CCATGCCCGG
35751 ACAATTTTTT TTTTTTTGAG ATGGAGCCTT GCTTTGTCGC CCAGGCTGGA
35801 GTGCCAGTGG TGCAATCTTG GCTCGCTGCA GCCTCCGCCT CCCAGGTTCA
35851 AGCAGTTCTC TTGCCTCAGC CTCCTGAGTA GCTGGGATTA CAGGCGTGCG
35901 CCACCACACC CTGCTAATTT TTTGTATTTT TAGTAGAGAC AGAGTTTCAC
35951 CATGTTGGTT AGGCTGGTCT CGAACTCCTG ACCTCGTGAT CTGCCTGACT
36001 CGGCTTCCCA AAGTGCTGGG ATTACAGGCA TGAGCCACTG TGCCCAGCCT
36051 TCCGATAATT TTTGTATTTT TCGTAGAGAT GGGATTTCGC CATGTTGGCC
36101 AGGCTGGTCT CAAACTCCTT ACCTCAAGTG ATCCACCCGT CTTGGCCTCC
36151 CAAAGTGCTG GGATTACAGG CGTGAGCCAC CACGCCTGGG TTTTTGAACA
36201 TTTTTAAGAA GCTTACCATT TTTTCGAAAT AGCTAGTTCC ATTTTACACA
36251 TAACTTCAGC TAGGCATGTT GCCTCATGCC TGTAATCCCA GCACTTTGGG
36301 AGGCCGAGGT CAGAGAGTCA CTTGAGGCCA GGAGTCAACA TAGCTCCTGT
36351 GACCAGCCTG GTCAACATAG AGACTCTATC TCTACCAAAA AAAAAAAAAA
36401 AAAAGTAAC CAGGTGTGGT GGTCCATGCC TGTAGTCCTA GCTCCCCAGG
36451 AGACTGAGGT GGGAGGAATG TTTGAGCCCA GGACTTCAAG GCTGCAGTGA
36501 GGCAAGATTG CACCATTGCA CCCCAGCTTT GGGGACAGAG TGAGAGACCC
36551 TGTCTCAAAA ACAAAATAAG GCTGGGCGCA GTGGCTGTCC GGGCGTCGTG
36601 GTTCACGCTT ATAGTCCTAG CACTTTGGGA GGCCAAGGTG GGCAGATTGC
36651 CTGAGCTCAG GAGGTCTAAG ACCAGCCTGA GCAACATGGC GAAACCTCAT
36701 CTTTGCAAAA CATACAGAAA AAAACAAAAA AAACCACAAA ACCTCTAGTT
36751 GCCAGTTATT TTTTTTATTT ATTCCTAGTG ATTCTTCTTT TTTTCTTTTT
36801 TCTGAGACAA AAATTTCACT TTGTCTCCCT CGCTAGAGTG CAGCGGTCAG
36851 CTCACTACAT GATTCTTTTA GAGACATGTT AATTCTTTAT ATTGAGCTGA
36901 AGCCTGTTTC TTTTACTTCT GTCTCTTCTT ATTCCTCCGC CTTGTAGAGC
36951 TGCCTGAATC AGATTAATTC CTCTTTTATT GGCAAGCCTG CCCTTCAGAT
37001 TGATCTTATC ACAACCTTTC TTCTACCTCT GAAGTCCTCA TTCTTTCCTG
37051 TAATGATATT TTCAGAACCT TGTGCAATTT GGGTTATTCT TACATTTTAT
37101 AAATGCCTTT TATTAAATTT GATTTCTTAA ATCAAGTATG AGATATAACA
37151 CATGAGGTAA ATCCTGTCTT GATTTGGAGC CTGAATGAAT TTCTCTCTTG
37201 AACTTCAAGG GCTCATGGCC CTTTCTTATT ATTAATCAAA GACAACCATT
37251 TGTTGTTTCA GTAGCTATAT TATTTCTAGT TTGGGTCTTA AGGTTTTTGA
37301 TTTGCTTGTT TTTTCTTTTT TCTTTTTTTT TTTTTGAGA CGGAGTTTCG
37351 CTCTTGTTGC CCAGACTGGG AGTGCAATGG CGTGATCTCG GCTCACTGCA
37401 ACCTCCGCCT CCCAGGTTCA AGCGATTCTT CTGCCTCAGC CTCCCTAGTA
37451 GCAGGGATTA CAGGCATGTG CCACCACGCC GGGCTAATTT TGTATTTTTA
37501 GTAGAGATGG GGTTTCTCCA TGTTGGTCAC GCTGGTCTCG AACTCCCGAC
37551 CTCAGGTGAT CCGCCTGCCT TGGCCTCCCA AAGTGCTGGG ATTACAGTCG
37601 TGAGCCACGG CGCCTGGCCG ATTTGCTTGT TTTTAATTAA AATAGGGGCC
37651 TTGGCCAGGT GCAGTTGTTC ACCCCTGTAA TCCCAGTACT TGGGAGGCT
37701 GAGGCAGGCA GATCTCTTGA GTTCAGGAGT TCAAGACCAG TATGGGCAAC
37751 ATGGTGAAAC CCTGTCTCTA CCAAAAACAC AAAATTCAGC CAGGCATGGT
```

FIGURE 3, page 12 of 21

```
37801 GGTGTGTCCC TGTAGTTCAA GGTACTCAGG AGGCTGAGGT GGGAGGATTG
37851 CTTGAGCCCG GAGATGGAGG TTGCGGTGAG CCAAGATTGT GCCATTTGCA
37901 CTCTAGCCTG GGCAACAGAG CGAGACCTTG TTTCAAAAAA AAAAAAGAAG
37951 AGGGTCTCAC TTTACACTTC TGTGACTGGT GTTTTAAAAA TCTAAACACA
38001 GGCCGGGCAC GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCAGA
38051 GGCACGCAGA TCACAAGGTC AGGAGTTCGT GACCAGCCTG GCCAGCATGG
38101 TGAAGCCCAT CTCTACTAAA AATACAAAAA AATTAGCTGG GCATGGTGGC
38151 AGGTGCCTGT AATCCCAGCT ACTTGGGAGG CTGAGACAGG GGAATCACTT
38201 GAACCCAGGA GGCGGAGATT GCAGTGAGCC AAGATTGCGC CATTGCACTC
38251 CAGCCTGGTG ACAGAGCGAG ACTCCGTCTG AAAAAAAAAA AAAAAAATCT
38301 AAACACAAGA TTTTACTTTT AATCCTATCA TTTCCTCTTG CTTGGCTTCA
38351 GTAATCCTTC AAGTTTTCTA GGTCTTTTCA AAATCTTGAT TCTGTTGATT
38401 TATATTTTAA TTATCTTTTC CTTTCAGCTT TTCCTGTTCA GGTGTGACAT
38451 CTGGGTCTTT ATCTGAGTTT TATTAGATTA TAAAACATTC AGCAAGATAG
38501 GGCAGGTACT GAGTCCAGTT GTACACCATG GAAGGCCTCT TTCTGTGATT
38551 GTTCATTCAT GAGGCTTTAT GAAAATGTCT ACATTACACC AGGCACTTGG
38601 AGGTTACAGA GATGAATAAA ACATAGTCCA TTAGGAGGCA GACAATGGGA
38651 GAGACAAACA TGGGAAAAAG TTACTCTGAT TATGAGGAGT AATGAGAATT
38701 ACATATGAAG GAAAGTATTG TTAGTACTGT TAGGATTTAG TGTCAGGAAA
38751 GTTTTCAGAG TAGCAAGGAA ACATCAGAAA TTTTACTCTT TCTGCCAGGC
38801 ATGGTGCATG TATTATTCTG TTCTCACACT GCCACAAGGA ACTGACCAAA
38851 ACTGGGTGAT TTATTAAAAA AAAGGTTTAA TTGACTCATA GTTCTGCATG
38901 GCTGAGGAGG CCTCAGGAAA CTTACTGTGG CAGAAAGGGA AGCAGGCACG
38951 TCTTACATGG CAGGAGGCGA GAGAGTGTGA AGGAAGTGAA GGGGGAAGAG
39001 CCCCTTATGA GACCATCAGA TCTTGTGAGA ATTCATTCAC TATCACTCGA
39051 ATGGGGGAAA CCGTCGTCAT AATCCAATCA CTTCTCCATA ATCCAATCAC
39101 TTCCCTCAGT GATTACAACT TGAGATGAGA TTTGGGTGGG GACACAGAGC
39151 CAAACCATAT CAGTGCCTGT AGTCCCAGTT ACTTGGAGGC TGAGGCAGGA
39201 GGAACACTTG AGCCCAGGAG TTCAAGATCT GCCTGGGCAA CATAGCAATA
39251 CCTCCATTTT GGATAAAAAG GAAATTTTAC TTTTTGGGTG CCATTGCTTA
39301 GTTTAATCAG CTGTAACTTC TTGTTGACTT TTAGTCAAAA AACAATTTTT
39351 CCTTCTATCT TTGTGAAAGA GGTTGGTGAG CAAGGAAGAA AAGGAAACTT
39401 GCTTTATTGA GCAGCTTCTA TAGTCAGGCA CATTTTACAA ACATTAGTTC
39451 ATTTAAACCC CTTTAGCTGT TGTACAAGGT GAATGCTATC TAGCATTTAC
39501 AGATGAAGAA ACTGTTAGGT GACTCTCCCT AATATTAAAT AACCAGGAAC
39551 CTGGATTTGA TGTTTTGAAG TCAGGGTAGC TTGATCCTCG AGTTCATGCT
39601 TCCTCCAAGG ATACACTGAA AGACTTTGAG CCTCTTTTTT TTTTTTTCTC
39651 TTTTTTTGAG ACAGGATCTG GCTCTCTTGC CCAGAGTGCA GTGGTGTGAT
39701 CTCAGCTCAC TGCAACCTCT GCCTCCTGGG CTCAAGCGAT TCTGCCTCAG
39751 CCTCTCGAGT AGCTGGGACC ACAGGCGCAC GCCAGCATAC TTGGCTAATT
39801 TTTGGATTTT TAGTAGAGAC AGGGTTTCAC CATGTTGGTC AGGCTGGTCT
39851 CGAACTCCTG AGCTCGTAAT CCGCCCGTCT CGGCCCCACA AAGTGCTGGG
39901 ATTACAGGCG TGAGCCACCG ACCCAGTCCC AACAGTTTTT TAAAACCCAG
39951 AACTATAATG CAATAATGTT AGCATTTGTT TTGGGAGTTT GAGCCTAAAT
40001 GGTTGAAGTG CAGTAAATTG TTCTTAAAAT ACGTTTTATG AAAGTATTTG
40051 GAGTCTCTTC CTTACATTTT TTTCTCTAGC ATGAAGACAA CACCTAGCCA
40101 GGCATGGTGG CTCATGCCAG TAATGCCAGC ACTTTGGGAG AATGAGTTAG
40151 GATAATTGCT TGAGTCCAGG AATTTGAGAC CAGCCTGGGC AATGTAGCGA
40201 GACTCTGTCT CTACAAAAAA GAAAAAATTA GCCGGGTGTG GTGGCATGTG
40251 CCTGTAGTCC CAGCTACTCA GGAGGCTCAG GTGGAAGGAT TGCTTGAGGT
40301 GGGAGGTTGA GGCTGCAGCG AGCCATGATC ATGCCACTGT ACTCAGCCTG
40351 GATGACAGAA TGAGACGCTG CTTGAGAGGG GAAAAAAAAG ACACCTGCTT
40401 GGGATGATTA AAGTTCTGTC TTGACTGGTA GTTATTTGAA TTAGGTCCCT
40451 CCAGTGCTTT TAATCATGGT AGAATGTGCT AGCAAGTGAG TTTGTCTTAC
40501 ATGGAAGAGT TCTGTGTTCA AGGGCTTTCG GCCAGTGGCA TTCCTAAACA
40551 CAGTGTTAAA GGCGGTAGGG AATGTGAAAA GTATGACATA GTTCCTGCTC
40601 TCAACAGCTT GTAATTTTAG TATTATTATC GTAAGCTCAA TTGTAGGTAC
40651 TACTTCTTTT CTGGACTTTC AGGTGCTTAT TACCGTGCAA TTTAGTGGTA
40701 TGAGTTGAGG ACTAATGTTT CTATATCACA TCCTGATAAT CTCCACAGTT
40751 ATGAAACTA AACTATTTCC CCTCCCTCCT ACACTTTTCC CCAACTTTAT
40801 TTTAATGGAA TTGTTTGGAT TTCTTGATTG TTTTGTAATA GTGGGACACA
40851 GCAGGCCAGG AAAGATTTCG AACAATCACC TCCAGTTATT ACAGAGGAGC
40901 CCATGGCATC ATAGTTGTGT ATGATGTGAC AGATCAGGTA AGTTCCAAGA
```

```
40951 GGAGATTGTG TTACAGTGAC CAAGTAGGAA GCCATTATTT GATTAATGTC
41001 AGATTCATTT ACTACTTCAT ATATAAGCCA TCAGTATTAA TTTTATGGCA
41051 GAAAACTTTG TCCACTCTCA AATATAAATG TGAATCACTT AAAAGACATT
41101 TGTTTTCCTG TAATAAATAA AAGATTAGTA ATTAGTTTTA CGTTTGCTTT
41151 CAAGGGATTC TGGTTGTATT TATTGTCAAC TAAATAACTT TGATCAAATA
41201 GCCAAGACTC TAACATATAG GCAAGAGTTT GTAGGGAATC GTGAGTTGCT
41251 TGGCTTATAC TGTGTTCTTG GTGTTAAGTA TTAACAGGAA TATGGCCTGG
41301 TAATTAGAAC TTGTCCATCA GAATTGCCAA AAGTGGGATT CGGGGGTCTC
41351 TGCCTATGGA GGATGTGGTT CAGAAATAAA GAATTTGAAT AGGATAAGCT
41401 GTAGGAGGAT CTTAGTATGA GAATGAGTAT CTGAAGATTA GCTGTGAGAG
41451 AGGGCAGAGC GATGGAGGGA ACAATGTGGG ACAGTGTGAA GCATGTGATC
41501 CAGGGGCCAT AACTTTTTTT GTTACTATTT TTTTAAATCA GAAACTTAGA
41551 TTTCAGTGTC CTTTCTATCA AAGAAAAGGA CAAAAGATAA ACGTTCAAAA
41601 TTGGAATTTA TTTTTCTTTT GGCAAATGTT AAATCTCACC TCTAATGAGA
41651 AATCATAGCT AATTAGGAGA TAACTTACAT GTAAGCATTT AGATTCAGTG
41701 CCATTAGAAG TGCTGGGTGG GTGATATCTG CAGGAGAAAA AAATGATGCT
41751 AGTTTAAAAA ATCTCTACTA TTACCGTGAA ATATTTTTAA ATGAAAACTT
41801 TCGTCCTCTA AATATGACTG TGGAAAAGAA AATGAGTATA TTTAATAACA
41851 TCTTTTGACA TCTCTAGTAG TAACAGTAGG TCATCTTATT CATAAACCAA
41901 AATTTTACCA AATTTCAGGC CAGGCGCAGT GGCTCATGCC TGTAATCCCA
41951 GAACTTTGGG AGGCCGAGGC GGGCGGATCA CCTGAGGTCA GGAGTTAGAG
42001 ACTAGCCTCG CCAACATGGC AAAATCCCAT CTCTAGTAAA AATACAAAAA
42051 TTAGCCAGGC GTGGGGGCCC GTGCCTGTAA TCCTAGCCAC TTGGGAGGCT
42101 GAGACAGGAG AATCGCTTGA ACCCAGCGGG CAGAGGTTGC AGTGAGCCGA
42151 GATCGCGCCA TTGCACTCCA GCCTGGATGA CAGAACAAGA CTTTGTCTCA
42201 AAAAAAAAAA AAAAAAAAAA AAAAAAATTA ATCAAATTTC AAAACCAGGT
42251 TTTGTAGTAC ATTTAAATTG CATATTCCAA AGCAGTTGGG TTTGCCTGCG
42301 TTGCAGTTTA ATATTAAGCT ATACTTCCCT TTCAAATAAG GTATTTTCAT
42351 CGTTAAGCCT GTAAATTCTA GTTTGTCATT GTTTAGATAT TTATAGTCAT
42401 TTTAATATAT CTGTTTACGG CCAGCTGCAA TGGCTAACAC CTGTAAACTC
42451 AGCACTTTTT GAGGCCAAGG TGGGCCGATT GAGCTCAGGA GTTCGAGACC
42501 AGCCTGGGCA ACATAGTGAA ACTCCATCTA TACAAAAAAT CCAAAAAAAA
42551 AAAGACAGGT GTGGTGGCAT GTGCCTGTAG TCCCAGCTAT CCCGGAGGCG
42601 GAGGCGGGAG GATGGCTTGA GCTTGGGAGG TCGAGGGTGC AGTGAGCTGT
42651 GATTGTGCCA CTGCACTCCG GCCTAGGTGA CAGAGCAAGA CCCTGTCTCA
42701 AAAAAAAAAA TCTCTTCACT CCTTAGCAGT GGTTATTTTG TAGCTAGAGT
42751 TGTCTCACTA GCTCTTTGTT ATTTGTCTGT TAGGTCAGGA ACGATGTTTC
42801 TGTTTATTCC AGAACTATAT TATCGAACTA TATTATCAGT CTTTCAAATG
42851 TCTTTTTAGG AGTCCTTCAA TAATGTTAAA CAGTGGCTGC AGGAAATAGA
42901 TCGTTATGCC AGTGAAAATG TCAACAAATT GTTGGTAGGG AACAAATGTG
42951 ATCTGACCAC AAAGAAAGTA GTAGACTACA CAACAGCGAA GGTATGTTTA
43001 AAGTTTAATT TTCATACTGA ATTTGAAGGT GTTGAATTAT GTATGGGTTC
43051 TGCAGTAACA GTAAGGCCAC AGCCTTTTAA AAATATGTGC ACTAGAATAC
43101 TGTGACAGTG ACAATTTGTG TAGCATCTGT TTGGATCCAA TGAACTTAGT
43151 TCCTCACGCT CCATTATGGA TGGTAGAAAT GCAGTAAGAA TTAGTGAAAA
43201 AGATTTTTCA GTGTTAATTG TGCCTCATTA TTCTCTTAGG AATTTGCTGA
43251 TTCCCTTGGA ATTCCGTTTT TGGAAACCAG TGCTAAGAAT GCAACGAATG
43301 TAGAACAGTC TTTCATGACG ATGGCAGCTG AGATTAAAAA GCGAATGGGT
43351 CCCGGAGCAA CAGCTGGTGG TGCTGAGAAG TCCAATGTTA AAATTCAGAG
43401 CACTCCAGTC AAGCAGTCAG GTGGAGGTTG CTGCTAAAAT TTGCCTCCAT
43451 CCTTTTCTCA CAGCAATGAA TTTGCAATCT GAACCCAAGT GAAAAAACAA
43501 AATTGCCTGA ATTGTACTGT ATGTAGCTGC ACTACAACAG ATTCTTACCG
43551 TCTCCACAAA GGTCAGAGAT TGTAAATGGT CAATACTGAC TTTTTTTTTA
43601 TTCCCTTGAC TCAAGACAGC TAACTTCATT TTCAGAACTG TTTTAAACCT
43651 TTGTGTGCTG GTTTATAAAA TAATGTGTGT AATCCTTGTT GCTTTCCTGA
43701 TACCAGACTG TTTCCCGTGG TTGGTTAGAA TATATTTTGT TTTGATGTTT
43751 ATATTGGCAT GTTTAGATGT CAGGTTTAGT CTTCTGAAGA TGAAGTTCAG
43801 CCATTTTGTA TCAAACAGCA CAAGCAGTGT CTGTCACTTT CCATGCATAA
43851 AGTTTAGTGA GATGTTATAT GTAAGATCTG ATTTGCTAGT TCTTCCTTGT
43901 AGAGTTATAA ATGGAAAGAT TACACTATCT GATTAATAGT TTCTTCATAC
43951 TCTGCATATA ATTTGTGGCT GCAGAATATT GTAATTTGTT GCACACTATG
44001 TAACAAAACA ACTGAAGATA TGTTTAATAA ATATTGTACT TATTGGAAGT
44051 AATATCAAAC TGTATGGTGA TAAGTATTGT TTTGATTCTT ATGGTTAAAG
```

FIGURE 3, page 14 of 21

```
44101 GGAAATAGAG CCTTGCATTA TATTCAACAC AGCCATTTGT GTGTGCACAA
44151 TGCAAACTAA GGTATTCTAG ACCTATCTTA GAGCAGCATC CAGTATTTGC
44201 TTTCTAGATA ATATGCCCAA TAACATGACC TAGAGGGGCT TCTGTGCTGT
44251 GTAGGGATTT AACCAACTTC AGTGGTTCAG GGAGCTCAAA CTATATGTAA
44301 AACAAGTTTA GAATGTATGC TATCTAGCCC GTTATCTCTG ATCCTTCTCT
44351 AAAACCATTT GAAATAGCTT CATTGATCAA CATTTCATAA ATGCATCTGT
44401 GGTAGAGGTA GAAAGCAGCA CCTTTCCTAA TTGGCAAATG ATCAGACTAA
44451 TGTGTGCTAA TGTTTTTCTT CCATGCTTTC AGTCAGATTC AACTATTTTA
44501 TCCTCCACAG TTGCTTAACT TGGTGTTGGA GGAGGGTTTA AGCATTAAGA
44551 TAGGAAGCAG GAAATTTGAT TGCTCTAAAT TTAGAAATTA TATCCCTAAA
44601 AATTAAAACA TGAATACTGG GTGGTAATGA TAATTGAGGC AAATGTATTT
44651 ATTTTGGTGA CATTTTGCAT ATATGAAGAT TTTCTGAAAT AGGACCTTCA
44701 AGATCCTAGG GGGTTTTGTT TGGTTTTTAA TTGTGAGGAA TAAAAAATCT
44751 TCTGCCCACA CTGGCATTTT AAGGTGACTG AGGTCAAACG TTGTTTCCTT
44801 AGGTTGAAAT AGCAGCCAAA ACATTCTTCA CGCAGGGGCT TGGGATATGG
44851 CTGCTGGCAA CACATTTTGT TGTGGGCTCC TTAATTTAAT GATAAAATTT
44901 AAGCTAAACA CAAGCCAAAA ATGAATAGGT TTTTTTAATT TTTATTTTTC
44951 ACTAAACAGG CAATTGAAAT ACATGGTACA AAAATAAGTG GTAAGATAAT
45001 TGTAAAATGA AATGGACAGA ATATTCAATT TTCCATCTAT GAAAATTTCA
45051 CAATAAAAAT CATAGTTTAC TTTGTATTAT AGGCGTGCTT GGTGGATCTA
45101 TTCATCCTCA CATAAGGCAA CTGACAAATT CCTGAAGTTA CCAATAGTTA
45151 TTTTGGTGAA GATCTTTAAT GCTTCAGAAG TTTTGTTTTT GCCTTAATAC
45201 AGTATAAAGG GGGAAAGAGT TCAGAAACTA TTTTCTAAAG TAGCTAAATG
45251 ACACAAAACA AATGTCAAGA TACTGTGATG CCATGCCGTG CACTTCATTT
45301 TTACACAGTA AAAGTTGTTT AAATTGTCAG CTTATTCTTG GTGAGTTAGC
45351 GGAAACATTA CATGAACTTA AGATGAGCAT ATTTACAGAC TTAAGTTTGG
45401 AAAATTCCAG CGTTCTTTTC CCCATGGCAG TAAAGATTGG GATTTACAAC
45451 AAATTTCAGC ATGCCTTAAG ATTTGCTTCT ATGTATACGC CAATAAATGT
45501 GGTTCTGGAA AAAATATATA CCCCTTTATA CCCCCATTTT CAAGTACAAA
45551 CGGTTCAAAG CTACTACAGG TTTTAATAAT CTGTTCACTT AGTAAAGGGA
45601 ATTACCACTT GTTCTAAATA TAAGGTGCTG CCATAAATTA GTTTACATAG
45651 TGAAGAAGAG TGTTCTTAAA TCTAAGCAGC TGCACACTCT GTGAAATCCT
45701 TTCAGAATGA TAGTCATTGT GGTCTGAGCA GTAATTTCCT ATTCTTCGAC
45751 CTTGGATTGA ATTTCCCTTA GCCTACATCT TGCCTTTCCA GCATATCTTA
45801 CCTCAAACCT TCTTTGTGTT CCATTCCCAC CTAAGCTTCA AAATAGCCCT
45851 GTGTTGACGT CGTCTTCCAT TTGCTGAGCT TACCTATGGA TCTCCAAGAA
45901 CCCAGATCTT GAAACTGCTG ATCCAGCTTT GAGTATCATC ACTTCCCTGT
45951 GGATTTAACT TCCATTAATT TTAAGGGACT ACTAAGTTAT TCCAGTGTGG
46001 CATCACAGTG CAGTTAGCAA GCTCAGCTAC TTGACTCTAA TTTGGCCATG    (SEQ ID NO:3)
```

FEATURES:

```
Start:    2181
Exon:     2181-2203
Intron:   2204-27090
Exon:     27091-27163
Intron:   27164-33853
Exon:     33854-33949
Intron:   33950-42859
Exon:     42860-42991
Intron:   42992-43239
Exon:     43240-43434
Stop:     43435
```

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
397	T	--	Beyond ORF(5')			
2326	A	G	Intron			

```
3486      C         A         Intron
6651      -         A         Intron
8190      T         -         Intron
8281      T         C         Intron
11546     A         G         Intron
11670     C         T         Intron
11688     A         G         Intron
14938     A         C         Intron
22261     G         A         Intron
22852     G         A         Intron
27253     A         C         Intron
28098     -         A         Intron
28597     G         T         Intron
31431     C         T G       Intron
35704     C         T         Intron
35728     C         T         Intron
36690     C         T         Intron
41002     G         C         Intron
41033     A         G         Intron
43161     C         T         Intron
43765     A         G         Beyond ORF(3')
44713     G         T         Beyond ORF(3')
44831     C         T         Beyond ORF(3')
```

Context:

DNA
Position
397
                    TGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCCTCTCGGCCCACTGTAGCCTCCGCCTCCC
                    GGGTTCAAGCAATTTTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCGCCA
                    CCATGCCTGGCTAATTTTTGTATTTTTAGTAGAGACAGTGTTTCACCATGTTGGCCAGGC
                    TGGTCTTGAATTCCTGACCTCGTGATCTGTCCGTTTTGGCCTCTCAAATTCCTGAGATTA
                    CAGGCATGAGCCACCGAGCCTGGCCAGTTTTCTGAGTTTTTATTTGAAATCAAAATAAGC
                    [T,-]
                    TTTTTTTTTTTTTTTAATGGGCTTTAGAGTCCAGGGTAACGAACACTTTTTGGTGCCTATT
                    ACTGAACCATTCAGGGTATTCCTGGGGTGGTGACCGTGTTCATTTCAGAAACCAACATGT
                    TCATTTCAGAAACCAAACTCGGGTAACTTTTGATAAGTTCATCAACTAAGGCCCATGGCA
                    GAATTTGAGGGCTAAGGGGTGTAATTAGTGTATGGGTAGAAATAAGTGCCTTCTTTCTAT
                    ATTTTGGCGTTGTAGGAATTTAAAGTGATTCTGCAGTAAGTCTCAGGAGACAATTTTCTT

2326
                    GCTGATTGTGTTCTAGGGGACGGAGTAGGGGAAGACGTTTGCTCTCCCGGAACAGCCTAT
                    CTCATTCCTTTCTTTCGATTACCCGTGGCGCGGAGAGTCAGGGCGGCGGCTGCGGCAGCA
                    AGGGCGGCGGTGGCGGCGGCGGCAGCTGCAGTGACATGTCCAGCATGAATCCCGAATAGT
                    GAGTTCAGGAGAGCACCGGTCGGCTGGGTCCGTGGGCCAGCTTGGGGGATCTTAAAGGGG
                    TCGAGGAGGGTTGGGGCAGAAGTCGGGGCATCGGCTGGGGTGAGGCGAGGGTGATGGGTC
                    [A,G]
                    GGAGAGGCTGGCGGCCGGGAGTCGGGCCCCATTGTCTGACGCGGAGGGGCGGCCGCGCGG
                    GGGAGGGGTCGGGCCGGAGGGGTGAGCCGCCCGGGCCTGGACCGGGTCAGGTTAGAGGGC
                    CTGACTGCGGGGCGGGTGCTGAGGAAGCCTGCCGAGGGGCCTGGGGCGGTGTGAAGGGGT
                    ATCTTCTCTCGGAGGCAGTGACTTTTGAAGGAGGACTTGTCTCTAAGGGGAGGGGATGGG
                    GTGGGAGAGCCCTTCTAGAGGGCACTGTCAGACCCTGCGCCCGCACTCTGCGGAGCTGTC

3486
                    CTGGGAACTGGTGTTCACTTCCCTTGGGTAGAGTTTGTTGGGCTCTCCTCAATGGCCCTT
                    TAAAAATTTCCTCTACAGTTTACATGCATGTAAAGTAATGAATAATTGGAAGAGACCGAA
                    TTGGTATTCCTTTTCAGTGTCAAAGGCCTTTGAGGGATGGGGGAAAATCAGTATTTGTTG
                    TAAAAGTTGAGTTTATTTGCTGGTTTGGTCAATTACTGCTAGACATTTTCCCCTAAAAGG
                    TCCACCCACCAGTTTAGCTGACTGTCATATGTGTGTCACATGGCTCTTGCAAAATGCTTA
                    [C,A]
                    AAGTTTTGTAATAGTGTGGCTTGAAGCTGAAATCTTTTGCACTAAACAGAAACCGTAGTA
                    TTTTATTAGAATTTCATGCTTTAGAAGTTGAGGGTAGTGTTCTTGTAGTGACATTTGCTG
                    TGTTGACAGTTTAAAAAAATTTTTTTTTCAAGGGCTCCAAGGACAAAGTTGGTTTTGCAC
                    AGTTGAACGGAGGTGAACTTGAGGTTCTTAATTTAGTAGTTTTCTTGGTAACAATAAAGA
                    ACATGGATTTACTGCTTTATCGAGGTTTATAGACCTCTACTGTTCAGGAAATTTTCTGAA
```

6651    TTTCAGCACATTAAGAAATGCTTAACATGGCCAGGCGCAGTGGCTCACGCCTGTAATTCT
CAGCACTTTGGGAGGCCGAGGTGGGCGGATCATTTGAGGTCATGACCAGCCTGGCCAACA
TGATGAGACACTGCCTCTACTAAAAATACAAAAATTAGCTGGGTGTGGTGGTGCACGCCT
GTAATTCCAGCTACTCAGGAACCTGAGGCAGGAGAGTCACTTGAACCTGGGAGGCGGAGG
CTGCAGTGAGTCCAGATCATGCCACTGCACTCCAGCCTGAGGGACAGAGTGAGACTCCTC
[-,A]
AAAAAAAAAAAAAAAAAAAGAAAGAAATACTTAACATTATTCTCGTGATTATTCTCATAAC
ATTTTTCATAATCCACTGGCTTCCAGTGGATTTTTTTAGTGTCAAGAAAATAATTTTGAT
TGGTTCATCTTTAAGGAATGTGTTAAGAATAAAGCATGTCTACCTGTCTTCAGTATACCA
GCTAACTATAGTAGGAAGAAATATAGTAGTCTACTTAGATCAACTATAATTCTTTAATGC
AGAAAAAGTTTAAAGTATTTACCTTATTTTTAGCCCCCATCCCCTTAAGTATATCATGGC

8190    AGACCGGCCTGGCCAATGTGGTGAAACCCTGCCTCTACTAAAAACACCAAATTAGCTAGG
CGTGGTGGTGTGCGCTTGTAGTCCCAAGCTACTGAGGAGGCTGAGACAAGAGAATCGCTT
GAATCTGGGAAAAGAGGTTGCCGTGAGCCAAGATTGGCCACTGCACTCCAGCCTGGGTG
ACAGAGTGAGATTCTGTCTCAAAAAAATAAAAAATAAAAATTTCCCCCTTTAATCAAATT
AAGTTAAAATGAGGGATGTTAGACAGTTTTTAACCATCAAATATTTTAGTTTAGTTTTTT
[T,-]
TTTTTAACGTTGTCTTAAAGATGGAAGTGCTTCAAAATCAAATCTTCCTTGCCAGTTCTC
TACTTGGCTTCTTTTTTTTTCTTTTTGAGATAGAGTCTCACTTTGTCACTGGAGTGCGTT
GGCGTGATCTCGGCTCACTGCAACCTCCGCCTTCCAGGTTTAAGTGATTCTTCCACCTCA
GCCTCTCAAGTAGCTGGGAGTACAGGTGTGTGCCACCACACCCGGCTAATTTTTGTAGTT
TTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGCCTCAAACTCCTGACCTCGTGA

8281    CTGAGGAGGCTGAGACAAGAGAATCGCTTGAATCTGGGAAAAGAGGTTGCCGTGAGCCA
AGATTGGCCACTGCACTCCAGCCTGGGTGACAGAGTGAGATTCTGTCTCAAAAAAAATAAA
AAATAAAAATTTCCCCCTTTAATCAAATTAAGTTAAAATGAGGGATGTTAGACAGTTTTTT
AACCATCAAATATTTTAGTTTAGTTTTTTTTTTTTAACGTTGTCTTAAAGATGGAAGTGC
TTCAAAATCAAATCTTCCTTGCCAGTTCTCTACTTGGCTTCTTTTTTTTTCTTTTTGAGA
[T,C]
AGAGTCTCACTTTGTCACTGGAGTGCGTTGGCGTGATCTCGGCTCACTGCAACCTCCGCC
TTCCAGGTTTAAGTGATTCTTCCACCTCAGCCTCTCAAGTAGCTGGGAGTACAGGTGTGT
GCCACCACACCCGGCTAATTTTTGTAGTTTTAGTAGAGACAGGGTTTCACTATGTTGGCC
AGGCTGGCCTCAAACTCCTGACCTCGTGATCCACCCACCTCAGCCAAATTGCTGGGATTA
CTTGTGTGAGCCACGCGCCTGGCTTCTACTTGGCTTTTAAAGGGAATTTTGCTTTCTGAG

11546   GTTACATTTAACCCATTTATGGTCGTGTAGCCATACTCACGTTACATTTGATGCATCTGC
TCCCTTTGTGTCTATATACTCATATAACATTTTGCATAAAGTTATAGGCAGTTCACACCA
AGGCTGTTCATGAACCTCAGATTAAGAATACTTGATTTAGGAGATTGAAAACAGAAAAGA
GAATGTTAACTATCATTATCAATATTAAAATGTGAAAATCTGAGAGTGACAAAGCTTAGC
TTTAAATCTGGTATCCCAAACTCATTTGAGTTTTTTTTTTTTTTTTTTTTTTTGAGAC
[A,G]
AGGTGTCGCTTTGTCCCCCAGGCTGGAGTGTAGTGGTGTGATCTTGGCTCACTGCAACCT
CCACCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCTGAAGTTGCTGGGATTACAGG
CTGCGCCACCACGCCCAGCTAATTTTTTGTATTTATAGTAAAGACGGAGTTTCACCTTAT
TGGCCAGGCTGGTCTCAAACTCCTGATCTTGTGATCCTCCCGCCTCGGCCTCCCAAAGTG
CTGGGATTACAGGTGTGAGCCACTGTTCCCGGCCTAATTTGAGTTTTAAAATGTGGAGTT

11670   TGTTCATGAACCTCAGATTAAGAATACTTGATTTAGGAGATTGAAAACAGAAAAGAGAAT
GTTAACTATCATTATCAATATTAAAATGTGAAAATCTGAGAGTGACAAAGCTTAGCTTTA
AATCTGGTATCCCAAACTCATTTGAGTTTTTTTTTTTTTTTTTTTTTTTTGAGACAAGG
TGTCGCTTTGTCCCCCAGGCTGGAGTGTAGTGGTGTGATCTTGGCTCACTGCAACCTCCA
CCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCTGAAGTTGCTGGGATTACAGGCTG
[C,T]
GCCACCACGCCCAGCTAATTTTTTGTATTTATAGTAAAGACGGAGTTTCACCTTATTGGC
CAGGCTGGTCTCAAACTCCTGATCTTGTGATCCTCCCGCCTCGGCCTCCCAAAGTGCTGG
GATTACAGGTGTGAGCCACTGTTCCCGGCCTAATTTGAGTTTTAAAATGTGGAGTTTAAG
ATGTTAGTCTTAAAGTGGGTTAGATGAAATTTATAAAAATAGTCAAATAGCTAAATTTAT
AAAAGGCCATTTGAAACAATTTTGTGAAATATATAATGTGGATAATTATGTAGTGCTTTA

11688   TAAGAATACTTGATTTAGGAGATTGAAAACAGAAAAGAGAATGTTAACTATCATTATCAA
TATTAAAATGTGAAAATCTGAGAGTGACAAAGCTTAGCTTTAAATCTGGTATCCCAAACT

## FIGURE 3, page 17 of 21

```
         CATTTGAGTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAAGGTGTCGCTTTGTCCCCCAG
         GCTGGAGTGTAGTGGTGTGATCTTGGCTCACTGCAACCTCCACCTCCCAGGTTCAAGTGA
         TTCTCCTGCCTCAGCCTCTGAAGTTGCTGGGATTACAGGCTGCGCCACCACGCCCAGCTA
         [A,G]
         TTTTTGTATTTATAGTAAAGACGGAGTTTCACCTTATTGGCCAGGCTGGTCTCAAACTC
         CTGATCTTGTGATCCTCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCA
         CTGTTCCCGGCCTAATTTGAGTTTTAAAATGTGGAGTTTAAGATGTTAGTCTTAAAGTGG
         GTTAGATGAAATTTATAAAAATAGTCAAATAGCTAAATTTATAAAAGGCCATTTGAAACA
         ATTTTGTGAAATATATAATGTGGATAATTATGTAGTGCTTTATGTGTAGATTGGTGGTTA

14938    CATGGTAGTGTGCACCTGTAGTCCCAACCACTTGGGAGGCTGAGGTGGGAGGATTGCCTG
         AGGCCAGGAGTTTGAGACCTGGGCAGCATATGAAGACCCTGTCTCTAAAAAACTAAAAAT
         AAAAAATAGCCAGGTGTGGTTGGTGTGCTTGTGGTCCCAGCTACTCAAGAGGCTGAGGCA
         AGAGGGTTGCTTGAGCCCAGAAGTTGGAGGCTGCCGTGAACTGTGATTGCACCACTGCAC
         TTCAGCCTGGGTGACATAGCAAGACCCTGTCTCTGTGGTGGTGGTGGGTGGGGGTGGGGG
         [A,C]
         AGGGATTTAAGAAGGGTTTGTGAGGTATGTATTATTTATAAATGGGCTTTTAACTTTACC
         CTTCACATCTTGGGTTGAAATTAATTGTATCCATTCTCAGTTTTTCTGTCTTGCTATATA
         TTTAAACTTGGAGACTTAGAGGTCATGGATGTCTTTCTATGAAAAGCAAATGAAGCAGAG
         GGCTGCCTTCTCTTGCTGTAGAGGGCACACTTGCTGCAGAGCATGTTACTGTTTTATGCA
         TTGCTAGGCTTTGGGAGTTGTGACTTGTATGATCATAGTACTTACAACTATTAGTTGGCA

22261    CACCCACAGATAGCTATGTCAAACGTAAGGGTGGAGAAACACAGACCCCAAACTTCTCGA
         GGGTAGAAAATATGAGGTTATAGTAGATTAGAACTACAAAAAGCTAGAGGAAGTTCTGAA
         CTGGAAACAGTGGATAGGATTTACTAGAATAATTTACGAGGGTGACAATTGTAAATCTTC
         ATAGGTTTCTTTTTTTTCCTTTCTCTTTTTTTTTTTTGAGATGGAGTCTCGCTCTGTTG
         CCCAGGCTGGAGTGCAATGGCGCAGTCTCTCCTCACTGCAACCTCCGCCTCCTGGGTCCA
         [G,A]
         GTGATTCTCCTGCCTTAGCCACCCAAGTAGCTGGGATTACAGGCATCTGCCACCATGCTG
         AGCTAATTTTTGTATTTTTTTTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTG
         GTCTTGAACTCCTGACCTCAGGTAATCCACCCACCTTGGCCTCCCAAAGTGCTGGGATTA
         CAGGTGTGAGCCACCGCGCCCAGCCAAATTTTTATTGGTTTCTAAACTAGCGTAATTTAG
         TTTTTTTCACTTAAGTCAAAATTATATTATTGTAGGATAAAAACTTAGTGATCCAAATTC

22852    ATCCAAATTCATGAGGAATGAAGAATAAATACATTTAAAGTCTTACCATTTGCTAAATTA
         GTCTTGGCTCTTTGTACCAAAATTCTGTCCTTGTGCTCTGTAATTTTATATTTGTATATT
         TTCTATCAACATTTTTACTGTGTGGTGTTTTGTAAATTATAAAAACGTTTTAAAGCAAAC
         TCAGAACAATGAATTCTCACGAATATTCAGTATATTTACAGTTGAGAAATAAACTACTTC
         TGTAGTAGGTAATTTAAAATGTCCCAATGCAAGTTAACGTGTCACTGATCACGCTATTCA
         [G,A]
         GTGTGTGTCTTTGATAAGGGGAGGTGGGGAAGTTTGTGGGTTTGATTTTATTTGCCTTTC
         TCATGTGACTGTTGTCATGTTAGTAAACAAATGGTTTGCGAGAGAACCAGTAGTCTTTTG
         CAAAGATTGTCTTATACAGAGCACTCAATTCTTCATATTATTTATAATGGCTTTAATTTA
         AGCCTTAAATTATTAGAAACTCATAAATAATTTTTTTATTTGTTTTTTTGAGATGGAGTT
         TCGCCCTTATTGTCCAGGCTGAAGTACAATGATGTGATCTTGACTCACTGCAACCTCCGC

27253    GCTTAAGCCATGCATGGGCTTTATAGGAGATGTAGTCTTCACAGTGAGTTGTTATTTGTA
         GCTGTGTTTTTGTTTTTGTATAGCTTATAGCAATGCAGTGTGCTTTTTATTAACATCATT
         TTCTTTTTCTTTTTGCAGTGATTATTTATTCAAGTTACTTCTGATTGGCGACTCAGGGGT
         TGGAAAGTCTTGCCTTCTTCTTAGGTTTGCAGTAAGTTGAAATTGAAATGTCTTTACAAT
         TAATGGTACAATTAATGCTATGTATGTTTTCTAGGTAGATAAAATTAAACAGTTTTATTC
         [A,C]
         GAATAAGTTAATTCTTCCAGAATTTATATATTTAAAGACTCCAAATATACATCCCCAGTG
         GTATCTTGGACTGTTAAATAGAAAAATATTGTTGCTCTTAAAAGAAATTCAGTGAAGTCT
         GGTTATAAAGTCAGAATGTCTAATACTTTTGGTCAGAGTCAAACAGCAGTTCCAATATAG
         GCAGCAAGTTAAAGGGGTAGTTGGTGGCCTGTGTTGAAAGCGACTTGATGAAAATAAATC
         TTTAAATTAAACTTTAGTAGAATAAAAAGAAAAAGCAGAGCCAGGTGACGCAGTGGATCA

28098    CTTTAAATTTAGCATGTTTCCTGGCCAGGTGCGGTGGCTCACGCCTGTAATCCCAGCACT
         TTGGGAGGCCGAGACGGGCGGATCACAAGGTCAGGAGATTGAGACCATCCTGGCTAACAC
         GGTGAAACCCCGTCTCTACTAAAAATACAAAAAATCAGCTGGGTGTGGTGCCACACGCCT
         GTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGCGGAGG
         TTGCAGTGAGCTGAGATGGTGCCACTGCACTCCAGCCTGGCAACAGAGCAAGACTGTCTC
```

FIGURE 3, page 18 of 21

```
          [-,A]
          AAAAAAAAAGAAAAAAAATAAAAAAACAAATTAGCATGTTTCCCTTCTAGAGATCATTGT
          TTCTCAGAGCATGGACCAAAGACTCCTGGGGGTTACCAAGACCCTCTCAGGTAGCCCATG
          AGGTCAAAATATCCTAATAATACTAAGATGTTAGTATTTGTAAGGAAATATTTACTTGGT
          AATAATACTAATATAAAAGATGTTTGCGTTTTTCAGTGATGACATTGGCTCTGGTACAAA
          AGCATGTGGGTAAAATTGCTGCTGGCTTGGTACACATCAAGGCAGCGCTAAGCTCCAAAT

  28597   GATGTTTGCGTTTTTCAGTGATGACATTGGCTCTGGTACAAAAGCATGTGGGTAAAATTG
          CTGCTGGCTTGGTACACATCAAGGCAGCGCTAAGCTCCAAATTGTACTCATGGTGATGGC
          ATTCTTTACCTCTGTGCCCTCACAGGAACAAAAACAAGCCGTGCCATTTTTATTGAAGAT
          TGTCCTTGACAAAACAGTTAAAATGATTAATTTTTGAAAAATGTTGATCCATGAGTATTC
          CTTTAAAAATATTTGTGAAGAAATGGGAAGTTCACATAAAACAATGTTTTTTTTTTGTTT
          [G,T]
          TTTTTTTTTTTTTTTTTTGAGACAGATTCTGGCTGTGTTGCCAAGGCTAGAGTGCAGTGGC
          GTCTGGCTCCCAGGCTCAAGCTGTTCTCCCACTTCAGCCTCCCAAGTGGCTGGGACCTCC
          CAAGTGGATGCGCCATCATGCCTGGCTGATTTTTGTATTTTTTTGTAGTGACAAGGTCTC
          ACTGTGTTGCACAGGCTGGTCTCAAACTTCTGAGCTCAAGCGATGCATGTGCCTCAGCCT
          CCCAAAGTGCTGGAGAAAGCACTTTTTACTGCATACTGGCTAGTGTGTTGGTTATTTTGG

  31431   CTGCATTTTTTTTTTTTTTTTTGGTTTGAGATGGAGTCTCGCTCTGTCGCCCAGGCTGGA
          GTGCAGTCGTGCAATCTCGGCTCACTGCAGCCTCCACCTCATGGGTTCAAGCGATTCTCC
          ATCTTGGTCTCCTGACTAGCTAGGTTTACAGGCGTGTGCCATCACACCCACTAATTTTTT
          GTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGATC
          TAAAGTGAGCCTCCCACCTTGGCCTCCCAAAGTGCTGGGATTACATATGTGAGCCACTGC
          [C,T,G]
          CCTGGCCTCTATATACTTCTATAGTACCTGATACTTATTAGGCACTCAATTACAACATAA
          CTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGACATGCCTTGTCGCCTGGGCTGGAGTGC
          AGTGGCACAGTCTCGGCTCACTGCAACCTTCACCTCCCGGGTTCAAGTGATTCTCCTTCC
          TCAGCCTCCCGGGTAGCTGGGATTACAGGCGCCCGCCACCACGTCCAGCTAATTTTTTGT
          ATTTTTAATAGAGATGAGGTTTCACCATCTTGGCCAGGCTGATCTCAAACTCCTGACCTT

  35704   ATGTGTGATCATTGGTGTTTATAAGATTTGGGTGTGTATTCGTGTGTGAAACATTCATAT
          TTTGTTACTTTCCTGTGGCTGGAAGGGATCTTATAGGACACTGTCTTTCATCTTTGTCTG
          TCTTTCATCTTTAATAGGAATTTCTTTTCCATGCCTGAAGGCCTCATTTTGAACATTTTG
          TTTGTTTGTTTTTTTATTTTTTGAGATACAGTATTGCTCTGTCTCCCAGGCTGGAGTGCA
          GTGGCGCGATTTGAGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGTGATTCTCCTGCCT
          [C,T]
          AGCCTCCCTAATAGCTGGGATTACATGTGTGTACCACCATGCCCGGACAATTTTTTTTTT
          TTTGAGATGGAGCCTTGCTTTGTCGCCCAGGCTGGAGTGCCAGTGGTGCAATCTTGGCTC
          GCTGCAGCCTCCGCCTCCCAGGTTCAAGCAGTTCTCTTGCCTCAGCCTCCTGAGTAGCTG
          GGATTACAGGCGTGCGCCACCACACCCTGCTAATTTTTTGTATTTTTAGTAGAGACAGAG
          TTTCACCATGTTGGTTAGGCTGGTCTCGAACTCCTGACCTCGTGATCTGCCTGACTCGGC

  35728   GATTTGGGTGTGTATTCGTGTGTGAAACATTCATATTTTGTTACTTTCCTGTGGCTGGAA
          GGGATCTTATAGGACACTGTCTTTCATCTTTGTCTGTCTTTCATCTTTAATAGGAATTTC
          TTTTCCATGCCTGAAGGCCTCATTTTGAACATTTTGTTTGTTTGTTTTTTTATTTTTTGA
          GATACAGTATTGCTCTGTCTCCCAGGCTGGAGTGCAGTGGCGCGATTTGAGCTCACTGCA
          ACCTCCGCCTCCTGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCTAATAGCTGGGATTA
          [C,T]
          ATGTGTGTACCACCATGCCCGGACAATTTTTTTTTTTTTGAGATGGAGCCTTGCTTTGTC
          GCCCAGGCTGGAGTGCCAGTGGTGCAATCTTGGCTCGCTGCAGCCTCCGCCTCCCAGGTT
          CAAGCAGTTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGTGCGCCACCACA
          CCCTGCTAATTTTTTGTATTTTTAGTAGAGACAGAGTTTCACCATGTTGGTTAGGCTGGT
          CTCGAACTCCTGACCTCGTGATCTGCCTGACTCGGCTTCCCAAAGTGCTGGGATTACAGG

  36690   AAAAAAAAAAAAAAAAGTAACCAGGTGTGGTGGTCCATGCCTGTAGTCCTAGCTCCCCAG
          GAGACTGAGGTGGGAGGAATGTTTGAGCCCAGGACTTCAAGGCTGCAGTGAGGCAAGATT
          GCACCATTGCACCCCAGCTTTGGGGACAGAGTGAGAGACCCTGTCTCAAAAACAAAATAA
          GGCTGGGCGCAGTGGCTGTCCGGGCGTCGTGGTTCACGCTTATAGTCCTAGCACTTTGGG
          AGGCCAAGGTGGGCAGATTGCCTGAGCTCAGGAGGTCTAAGACCAGCCTGAGCAACATGG
          [C,T]
          GAAACCTCATCTTTGCAAAACATACAGAAAAAAAACAAAAAAAACCACAAAACCTCTAGTT
          GCCAGTTATTTTTTTTATTTATTCCTAGTGATTCTTCTTTTTTTCTTTTTTTCTGAGACAA
```

FIGURE 3, page 19 of 21

```
         AAATTTCACTTTGTCTCCCTCGCTAGAGTGCAGCGGTCAGCTCACTACATGATTCTTTTA
         GAGACATGTTAATTCTTTATATTGAGCTGAAGCCTGTTTCTTTTACTTCTGTCTCTTCTT
         ATTCCTCCGCCTTGTAGAGCTGCCTGAATCAGATTAATTCCTCTTTTATTGGCAAGCCTG

41002    GAGTTGAGGACTAATGTTTCTATATCACATCCTGATAATCTCCACAGTTATGAAAACTAA
         ACTATTTCCCCTCCCTCCTACACTTTTCCCCAACTTTATTTTAATGGAATTGTTTGGATT
         TCTTGATTGTTTTGTAATAGTGGGACACAGCAGGCCAGGAAAGATTTCGAACAATCACCT
         CCAGTTATTACAGAGGAGCCCATGGCATCATAGTTGTGTATGATGTGACAGATCAGGTAA
         GTTCCAAGAGGAGATTGTGTTACAGTGACCAAGTAGGAAGCCATTATTTGATTAATGTCA
         [G,C]
         ATTCATTTACTACTTCATATATAAGCCATCAGTATTAATTTTATGGCAGAAAACTTTGTC
         CACTCTCAAATATAAATGTGAATCACTTAAAAGACATTTGTTTTCCTGTAATAAATAAAA
         GATTAGTAATTAGTTTTACGTTTGCTTTCAAGGGATTCTGGTTGTATTTATTGTCAACTA
         AATAACTTTGATCAAATAGCCAAGACTCTAACATATAGGCAAGAGTTTGTAGGGAATCGT
         GAGTTGCTTGGCTTATACTGTGTTCTTGGTGTTAAGTATTAACAGGAATATGGCCTGGTA

41033    CTGATAATCTCCACAGTTATGAAAACTAAACTATTTCCCCTCCCTCCTACACTTTTCCCC
         AACTTTATTTTAATGGAATTGTTTGGATTTCTTGATTGTTTTGTAATAGTGGGACACAGC
         AGGCCAGGAAAGATTTCGAACAATCACCTCCAGTTATTACAGAGGAGCCCATGGCATCAT
         AGTTGTGTATGATGTGACAGATCAGGTAAGTTCCAAGAGGAGATTGTGTTACAGTGACCA
         AGTAGGAAGCCATTATTTGATTAATGTCAGATTCATTTACTACTTCATATATAAGCCATC
         [A,G]
         GTATTAATTTTATGGCAGAAAACTTTGTCCACTCTCAAATATAAATGTGAATCACTTAAA
         AGACATTTGTTTTCCTGTAATAAATAAAAGATTAGTAATTAGTTTTACGTTTGCTTTCAA
         GGGATTCTGGTTGTATTTATTGTCAACTAAATAACTTTGATCAAATAGCCAAGACTCTAA
         CATATAGGCAAGAGTTTGTAGGGAATCGTGAGTTGCTTGGCTTATACTGTGTTCTTGGTG
         TTAAGTATTAACAGGAATATGGCCTGGTAATTAGAACTTGTCCATCAGAATTGCCAAAAG

43161    AGTCCTTCAATAATGTTAAACAGTGGCTGCAGGAAATAGATCGTTATGCCAGTGAAAATG
         TCAACAAATTGTTGGTAGGGAACAAATGTGATCTGACCACAAAGAAAGTAGTAGACTACA
         CAACAGCGAAGGTATGTTTAAAGTTTAATTTTCATACTGAATTTGAAGGTGTTGAATTAT
         GTATGGGTTCTGCAGTAACAGTAAGGCCACAGCCTTTTAAAAATATGTGCACTAGAATAC
         TGTGACAGTGACAATTTGTGTAGCATCTGTTTGGATCCAATGAACTTAGTTCCTCACGCT
         [C,T]
         CATTATGGATGGTAGAAATGCAGTAAGAATTAGTGAAAAAGATTTTTCAGTGTTAATTGT
         GCCTCATTATTCTCTTAGGAATTTGCTGATTCCCTTGGAATTCCGTTTTTGGAAACCAGT
         GCTAAGAATGCAACGAATGTAGAACAGTCTTTCATGACGATGGCAGCTGAGATTAAAAAG
         CGAATGGGTCCCGGAGCAACAGCTGGTGGTGCTGAGAAGTCCAATGTTAAAATTCAGAGC
         ACTCCAGTCAAGCAGTCAGGTGGAGGTTGCTGCTAAAATTTGCCTCCATCCTTTTCTCAC

43765    AATGAATTTGCAATCTGAACCCAAGTGAAAAAACAAAATTGCCTGAATTGTACTGTATGT
         AGCTGCACTACAACAGATTCTTACCGTCTCCACAAAGGTCAGAGATTGTAAATGGTCAAT
         ACTGACTTTTTTTTTATTCCCTTGACTCAAGACAGCTAACTTCATTTTCAGAACTGTTTT
         AAACCTTTGTGTGCTGGTTTATAAAATAATGTGTGTAATCCTTGTTGCTTTCCTGATACC
         AGACTGTTTCCCGTGGTTGGTTAGAATATATTTTGTTTTGATGTTTATATTGGCATGTTT
         [A,G]
         GATGTCAGGTTTAGTCTTCTGAAGATGAAGTTCAGCCATTTTGTATCAAACAGCACAAGC
         AGTGTCTGTCACTTTCCATGCATAAAGTTTAGTGAGATGTTATATGTAAGATCTGATTTG
         CTAGTTCTTCCTTGTAGAGTTATAAATGGAAAGATTACACTATCTGATTAATAGTTTCTT
         CATACTCTGCATATAATTTGTGGCTGCAGAATATTGTAATTTGTTGCACACTATGTAACA
         AAACAACTGAAGATATGTTTAATAAATATTGTACTTATTGGAAGTAATATCAAACTGTAT

44713    AAGCAGCACCTTTCCTAATTGGCAAATGATCAGACTAATGTGTGCTAATGTTTTTCTTCC
         ATGCTTTCAGTCAGATTCAACTATTTTATCCTCCACAGTTGCTTAACTTGGTGTTGGAGG
         AGGGTTTAAGCATTAAGATAGGAAGCAGGAAATTTGATTGCTCTAAATTTAGAAATTATA
         TCCCTAAAAATTAAAACATGAATACTGGGTGGTAATGATAATTGAGGCAAATGTATTTAT
         TTTGGTGACATTTTGCATATATGAAGATTTTCTGAAATAGGACCTTCAAGATCCTAGGGG
         [G,T]
         TTTTGTTTGGTTTTTAATTGTGAGGAATAAAAAATCTTCTGCCCACACTGGCATTTTAAG
         GTGACTGAGGTCAAACGTTGTTTCCTTAGGTTGAAATAGCAGCCAAAACATTCTTCACGC
         AGGGGCTTGGGATATGGCTGCTGGCAACACATTTTGTTGTGGGCTCCTTAATTTAATGAT
         AAAATTTAAGCTAAACACAAGCCAAAAATGAATAGGTTTTTTTAATTTTTATTTTTCACT
         AAACAGGCAATTGAAATACATGGTACAAAAATAAGTGGTAAGATAATTGTAAAATGAAAT
```

FIGURE 3, page 20 of 21

44831   GGAGGGTTTAAGCATTAAGATAGGAAGCAGGAAATTTGATTGCTCTAAATTTAGAAATTA
TATCCCTAAAAATTAAAACATGAATACTGGGTGGTAATGATAATTGAGGCAAATGTATTT
ATTTTGGTGACATTTTGCATATATGAAGATTTTCTGAAATAGGACCTTCAAGATCCTAGG
GGGTTTTGTTTGGTTTTTAATTGTGAGGAATAAAAAAATCTTCTGCCCACACTGGCATTTT
AAGGTGACTGAGGTCAAACGTTGTTTCCTTAGGTTGAAATAGCAGCCAAAACATTCTTCA
[C,T]
GCAGGGGCTTGGGATATGGCTGCTGGCAACACATTTTGTTGTGGGCTCCTTAATTTAATG
ATAAAATTTAAGCTAAACACAAGCCAAAAATGAATAGGTTTTTTTAATTTTTATTTTTCA
CTAAACAGGCAATTGAAATACATGGTACAAAAATAAGTGGTAAGATAATTGTAAAATGAA
ATGGACAGAATATTCAATTTTCCATCTATGAAAATTTCACAATAAAAATCATAGTTTACT
TTGTATTATAGGCGTGCTTGGTGGATCTATTCATCCTCACATAAGGCAACTGACAAATTC